January 2015

QUANTITATIVE SAFETY ASSESSMENT OF AIR TRAFFIC CONTROL SYSTEMS THROUGH SYSTEM CONTROL CAPACITY

Jingjing Guo
Purdue University

Graduate School Form 30
Updated 1/15/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Jingjing Guo

Entitled

QUANTITATIVE SAFETY ASSESSMENT OF AIR TRAFFIC CONTROL SYSTEMS THROUGH SYSTME CONTROL CAPACITY

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

STEVEN LANDRY
Co-chair

KAREN MARAIS
Co-chair

CHARLES ROBERT KENLEY

BARRETT CALDWELL

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Steven Landry and Karen Marais

Approved by: Weyne Chen 12/11/2015

Head of the Departmental Graduate Program Date

QUANTITATIVE SAFETY ASSESSMENT OF AIR TRAFFIC CONTROL SYSTEMS THROUGH SYSTEM CONTROL CAPACITY

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Jingjing Guo

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2015

Purdue University

West Lafayette, Indiana

To my mother.

ACKNOWLEDGEMENTS

I first would like to thank my committee Dr. Steven Landry, Dr. Karen Marais, Dr. Charles Robert Kenley and Dr. Barrett Caldwell, for their valuable insights and thoughtful examinations of my dissertation and research work. Their expertise and suggestions were much needed and greatly appreciated in my investigation and compilation of this exciting research idea.

It has been a trying time pursuing this degree. I owe great debt to my advisors, Dr. Steven Landry and Dr. Karen Marais who took me on and never gave up on me. I was their faith in me and their generous help that guided me to the end of the tunnel. They are the best advisors I can ask for and the kinds I aspire to be.

I am grateful to my colleagues at the Landry lab and VRSS lab. Their help, support and friendship made it very easy to seek intellectual conversations, bounce research ideas and practice for exams.

Beyond Purdue, I was blessed with care and help from mentors, friends and families. Thanks to Dr. Larry Baxter, Dr. Corey Miller and Dr. Erica Carlson for being a friend, an inspiration and a role model to me – I desired a PhD because I thought their humbleness, intelligence, and kindness came with the degree. Thanks to Crystal, Richard, Amna, and Julian for adopting me in their diverse U.S. family and being my best friends. Thanks to

many other kind families and friends I met along the way. I would like to especially thank my real family for their unconditional love and support and often during the most difficult times. And finally thanks to Prof. Sveinn Olafaso for everything else.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Guo, Jingjing. Ph.D., Purdue University, December 2015. Quantitative Safety Assessment of Air Traffic Control Systems through System Control Capacity. Major Professor: Steven Landry, Karen Marais.

Quantitative Safety Assessments (QSA) are essential to safety benefit verifications and regulations of developmental changes in safety critical systems like the Air Traffic Control (ATC) systems. Effectiveness of the assessments is particularly desirable today in the safe implementation of revolutionary ATC overhauls like NextGen and SESAR. QSA of ATC systems are however challenged by system complexity and lack of accident data.

Extending from the idea “safety is a control problem” in the literature, this research proposes to assess system safety from the control perspective, through quantifying a system’s “control capacity”. A system’s safety performance correlates to this “control capacity” in the control of “safety critical processes”. To examine this idea in QSA of the ATC systems, a Control-capacity Based Safety Assessment Framework (CBSAF) is developed which includes two control capacity metrics and a procedural method. The two metrics are Probabilistic System Control-capacity (PSC) and Temporal System Control-capacity (TSC); each addresses an aspect of a system’s control capacity. And the procedural method consists fthree general stages: I) identification of safety critical

processes, II) development of system control models and III) evaluation of system control capacity.

The CBSAF was tested in two case studies. The first one assesses an en-route collision avoidance scenario and compares three hypothetical configurations. The CBSAF was able to capture the uncoordinated behavior between two means of control, as was observed in a historic midair collision accident. The second case study compares CBSAF with an existing risk based QSA method in assessing the safety benefits of introducing a runway incursion alert system. Similar conclusions are reached between the two methods, while the CBSAF has the advantage of simplicity and provides a new control-based perspective and interpretation to the assessments.

The case studies are intended to investigate the potential and demonstrate the utilities of CBSAF and are not intended for thorough studies of collision avoidance and runway incursions safety, which are extremely challenging problems. Further development and thorough validations are required to allow CBSAF to reach implementation phases, e.g. addressing the issues of limited scalability and subjectivity.

CHAPTER 1. INTRODUCTION

Safety assessment is the evaluation or estimation of the *nature*, *quality*, or *ability* of a system to maintain accident free operations. Safety assessments of safety critical systems such as the Air Traffic Control (ATC) systems paradoxically are extremely difficult, due to their complexity. Several researchers have proposed that viewing safety in these systems as a problem of control can be helpful. This research builds on that work, with the development of a concept of "control capacity".

1.1 Assess system safety from the control perspective

Safety assessments identify a system's vulnerabilities, termed hazards, its tolerance of hazards, as well as possible counter measures to handle hazards, in the presence of system disturbance, disruptions and degradation (Safe Work Australia, 2012).

The counter-hazard measures or safety measures can be viewed as safety controls, which maintain the system in a safe state, or return it to a safe state. Several researchers have taken this view, most notably Leveson, whose STAMP accident model is based on the idea that accidents happen when one or more of the control means to prevent or mitigate hazards is inadequate in some way (Leveson 2004).

Intuitively, it is clear that some systems are safer than others. It then follows that if safety is established through control, the "capacity" of this control can differentiate

different systems' safety performance and thus act as an indicator of each system's safety. I therefore propose "control capacity" as one potential way of assessing a system's safety. Most research taking the control approaches to safety has focused on using the concept to understand why accidents happen, or to guide design decisions, which are qualitative. In this research, I propose quantitatively assessing system safety from the control perspective via the quantification of this "control capacity".

1.2 Need for Quantitative Safety Assessments in ATC

The second motivation for this research is the need for comprehensive Quantitative Safety Assessment (QSA) methods in ATC systems. To address increasing traffic demands and aging infrastructures, air transportation systems around the world are undergoing revolutionary overhauls, represented by the Next Generation Air Transportation Systems (NextGen) and the Single European Single ATM (Air Traffic Management) Research (SESAR) (FAA, 2011, SESAR 2012, Brooker, 2008). NextGen for example will "redesign airspace and deploy new performance-based flight procedures, develop systems to help controllers better manage air traffic, and provide critical technologies and infrastructure for NextGen" (Scovel III and General, 2013). Implementation challenges of NextGen in the latest government reports share the themes of undefined benefits of NextGen and consequently ATC users' reluctance to invest on and adopt new components and procedures (Scovel III and General, 2013).

To verify and assess safety benefits of the new components or concepts of operations introduced to a system, all safety assessment approaches follow an underlying safety philosophy (whether explicitly formulated or not) (Brooker, 2002b). As shown in Figure

1-1, the expected level of safety must be determined, e.g. through modeling, prediction and validation. Only if the safety levels are not compromised when comparing the expected level of safety to a target level of safety, can the changes be approved. An essential part of making the comparison is to express and assess safety quantitatively, and thus quantitative safety assessments.

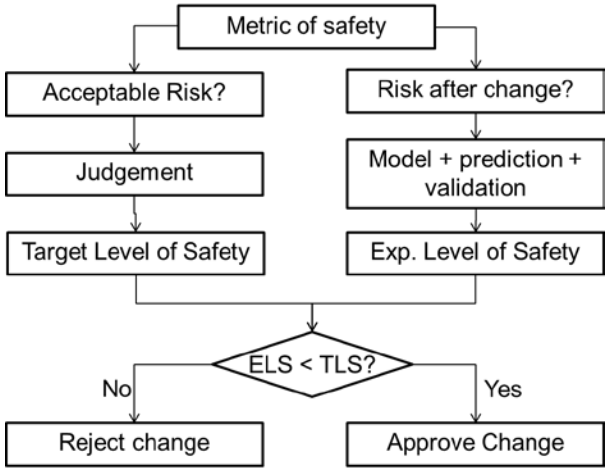


Figure 1-1 Safety Check of ATC Change adapted from (Brooker, 2002a)

1.2.1 Challenges faced by Quantitative Safety Assessment in ATC

Quantitative safety assessment in safety critical systems like the ATC system is extremely difficult. Two major challenges of making meaningful and credible quantitative assessments are 1) system complexity and 2) lack of accident data.

System complexity: The ATC system is a large scale, geographically distributed, and socio-technical system. It is extremely difficult to directly simulate the system's dynamics, or to accurately predict its future behaviors. To assure safety, safety critical systems in particular often have multiple redundant controls, which further increase the complexity of these systems. Another complication factor is the integration of human

components and automated systems. While such integration leverages the advantages of human and automation, it is a major source of uncertainty and further renders system's behaviors unpredictable (Wickens, 1998).

Lack of accident data: safety critical systems are engineered to be extremely safe. Air transportation in the United States currently has safety level of about 1 disastrous aircraft accident per 10^7 flight hours (≈1140 years) (Savage, 2013). The risk of a mid-air collision is estimated at $10^{-7}\sim10^{-12}$ per flight hour (Knecht, 1997, Blom et al., 2001).

Due to the rarity of air traffic accidents, to measure safety with statistical metrics as such, accident data needs to be collected over long periods of time. It does not tell if the system "*now* is safe, or if it is getting safer or less safe" (Brooker, 2007).

An alternative approach is to use incident data. The issue with incident data is its incompleteness due to many complication factors in the reporting of incidents (Reynard, 1986, Shorrock, 2005, Shorrock, 2007). For a more comprehensive survey of current safety metrics using incident data, see (Brooker, 2007).

Current QSA approaches are predominantly risk based. To assess a risk level of $10^{-7}\sim10^{-12}$ per flight hour for the ATC systems, many risk based QSA approaches are greatly challenged. The retrospective, data driven approaches must rely on very limited accident data that are collected over long durations, which are subject to significant fluctuations with occurrence of a single accident event among other drawbacks (Brook 2007). The prospective, simulation based approaches are faced with two major hurdles, 1) system complexity as discussed previously which yield results with magnitudes larger

uncertainties, and 2) very long simulation hours since one accident is expected to occur every millions of (simulated) flight hours (Blom et al., 2001).

The challenges to quantitatively assess the immediate safety performance of the ATC systems are not trivial; it requires years of data collection, exploration of many possible methods, and continuous validation and verification practices. The purpose of the proposal to use "control capacity" as a safety metric is a small yet necessary step towards exploring alternative safety metrics; it extends and builds upon the previous efforts to correlate system safety with system control. As is suggested by (Mannan, 2012), a diversity of metrics should be used to comprehensively describe the system's safety performance. A system's "Control capacity" is a relevant, but rarely explored addition to the current safety metric repository.

1.3 Research objectives and Scope

To reiterate the above discussions, the intended research objectives and questions are summarized in Table 1-1. The first research objective seeks formal definition of system control capacity, and establishment of its correlation to a system's safety performance. The second objective examines the viability of system control capacity in quantitative safety assessments of ATC systems. The intermediate research questions leading to the research objectives are listed in Table 1-1 as well.

Table 1-1: Research Objectives and Research Questions

Number	Objectives and Research Questions
Objective 1	**Correlate Control Capacity to System Safety Performance**
	Q1: What is system control capacity? Q2: How does system control relate to system safety? Q3: How does system control capacity relate to system safety performance?
Objective 2	**Examine Viability of Control Capacity as Safety Performance Metric in Quantitative Safety Assessment of ATC**
	Q1: How can system control capacity be quantified? Q2: How can the quantification of control capacity used for quantitative safety assessments? Q3: What are the implementation challenges using control capacity in ATC quantitative safety assessment? Q4: Can results of system control capacity based safety assessment be trusted?

1.4 Organization

The remainder of this document is organized as follows: CHAPTER 2 reviews previous research on ATC system accident prevention and safety assessment. CHAPTER 3 seeks formal definitions of system control capacity and details theoretic basis to use control capacity as a safety measure. CHAPTER 4 introduces two metrics for system control capacity: Probabilistic System Control Capacity (PSC) and Temporal System Control Capacity (TSC). CHAPTER 5 elaborates on the theoretic framework needed for quantitative safety assessments using the two proposed control capacity metrics. CHAPTER 6 demonstrates the use of the theoretic framework in a case of en route collision avoidance. CHAPTER 7 compares the theoretic framework with an existing

method in assessing safety benefits of a runway incursion alert system. CHAPTER 8 concludes the findings in the two case studies and about the method, and proposes directions for future work.

CHAPTER 2. SAFETY ASSESSMENTS OF AIR TRAFFIC CONTROL SYSTEMS

This chapter reviews researches and practices relevant to quantitative safety assessments of ATC systems. First, definitions and scopes of ATC accidents and safety concerns are reviewed. To assure safety, the previous and current approaches are categorized and summarized. As means to assess the effectiveness of safety assurances, both qualitative and quantitative assessment methods are then discussed. Qualitative safety assessments identify system hazards, which is often a prerequisite or a component for quantitative safety assessments. For quantitative safety assessments, the focus is on existing safety metrics. Finally, a brief review on the control related safety metrics is provided.

2.1 ATC accidents

2.1.1 Definitions

Safety and Accidents: Qualitatively, safety is the absence of accidents (Leveson, 2011). Accidents are events characterized by 1) loss or injury and 2) unknown times of occurrence. In different contexts, the former characteristic requires further specification, which can be arbitrary. For instance, the amount of loss that distinguishes “incidents” from “accidents” in aviation is specified by the International Civil Aviation Organization (ICAO) in The International Civil Aviation Annex 13.

Aviation accidents: In safety critical systems, accidents are often associated with disastrous consequences, e.g. loss of lives. The International Civil Aviation Annex 13 defines an aviation accident as “an occurrence associated with the operation of an aircraft, which takes place between the time any person boards the aircraft with the intention of flight until all such persons have disembarked, where a person is fatally or seriously injured, the aircraft sustains damage or structural failure or the aircraft is missing or is completely inaccessible” (Site, 1994). This definition specifies the general category of aviation accidents, as well as the severity of loss for an aviation mishap to be counted as an accident.

ATC accidents: Not all aviation accidents are ATC accidents; that is to say, not all aviation accidents are caused by ATC. For example, mid-air collisions and runway incursions are ATC accidents; structure failure and bird strike are not.

The ATC system is “principally a matter of preventing collisions with other aircraft, obstructions, and the ground; assisting aircraft in avoiding hazardous weather; assuring that aircraft do not operate in airspace where operations are prohibited; and assisting aircraft in distress” (Varon, 2000, FAA, 2015). This summary draws the boundaries of ATC’s “circle of influence”, beyond which the ATC has no control over.

2.1.2 Views of accidents

Understanding why and how accidents occur is necessary for eliciting lessons from past accidents and for accumulating knowledge of engineering safer systems in the future. Accident models are mental constructs that help humans to make sense of accidents, and to help develop counter measures to prevent future accidents.

Chain of events: The earliest and most popular accident model is the sequential events model, also known as the chain of events model. The chain of events model assumes that accidents occur as a results of series of hazardous events.

The first event sequence accident model, the domino model, was introduced by Heinrich in 1932. As shown in Figure 2-1, the accident outcome i.e. “injury” is at the end of a series of events: 1) social environment (cultural tolerance or incentive to risk taking), 2) fault of the person, 3) unsafe acts or conditions, 4) accident, and 5) injury.

The event chain is analogous to a falling line of dominos, hence the name dominos. In accidents, the occurrence of the leading event causes the next event to occur, and eventually the accident event at the end of the chain. The model also represents Heinrich’s belief that accidents are rooted in the deeper and broader environmental contexts, e.g. the social environment (Mannan, 2012).

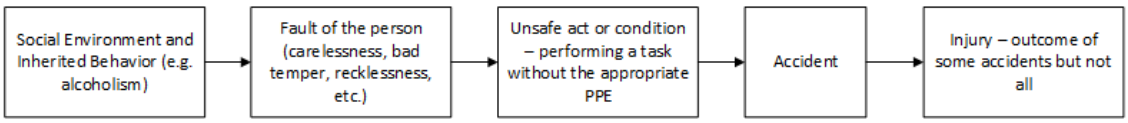


Figure 2-1 Heinrich's Domino accident model - a chain of events

Swiss cheese model (Reason, 2000)**:** A Swiss cheese model moves towards a system view that an accident is not simply a collection of mistakes, but inherent to the system structures and conditions within which it operates. The Swiss cheese model abstracts a system as layers of safeguards, analogous to Swiss cheese slices superimposed on each other. Each layer will have defects, like the holes on a Swiss cheese slice. Unlike actual holes on the cheese, "holes" in the system's safeguards open, close and shift over time.

Accidents happen when "holes" in the safeguards line up to allow an accident trajectory to penetrate through. Figure 2-2 is an illustration of the Swiss Cheese accident model adapted from (Reason, 2000) and (Maurino et al., 1995)

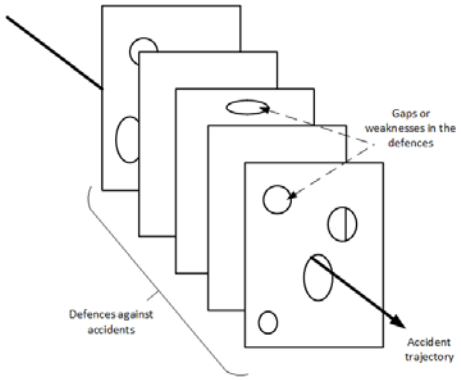


Figure 2-2 Swiss Cheese Model adapted from (Maurino et al., 1995)

If the chain of events interprets operation process of the system, then the Swiss cheese model focuses on explaining the system's role in accident occurrences and prevention. Instead of containing error events in the event chain leading to an accident, countermeasures derived from the Swiss cheese model center around making changes to operational conditions and system configurations, and strengthening defenses against accidents.

System-Theoretic Models (Rasmussen, 1997, Leveson, 2004): With the increase of complexity in modern engineering systems, e.g. socio-technical systems, sophisticated safety models are needed to represent and interpret the complex system dynamics that produce accidents. Stroeve et al. argue that accidents can be considered as an emergence of a system, a result of system behaviors and dynamics (Stroeve et al., 2009).

Rasmussen first promoted a “top-down” rather than “bottom up”, and a “functional abstraction” rather than “structure decomposition” approach to model accident causation. Rasmussen argues that risk management and accident prevention should be considered a control problem; it is a cross-disciplinary subject and requires all levels of society to be involved in this control process including legislators, managers, and operators (Rasmussen, 1997). A sketch of Rasmussen’s hierarchical system model is illustrated in Figure 2-3. In this model, the upper levels in the hierarchy have higher authority in decision making than lower levels and concern more strategic and long term goals, as opposed to tactical and short term ones, for the system. The decisions, in terms of policies, regulations, and judgments made at the upper levels, are passed down, concretized and actualized at the lower levels. Information on the effects of these decisions is continuously observed and reported back to the upper levels, in the form of data, reports, and reviews, which are then integrated to the upper levels to support future decision making activities.

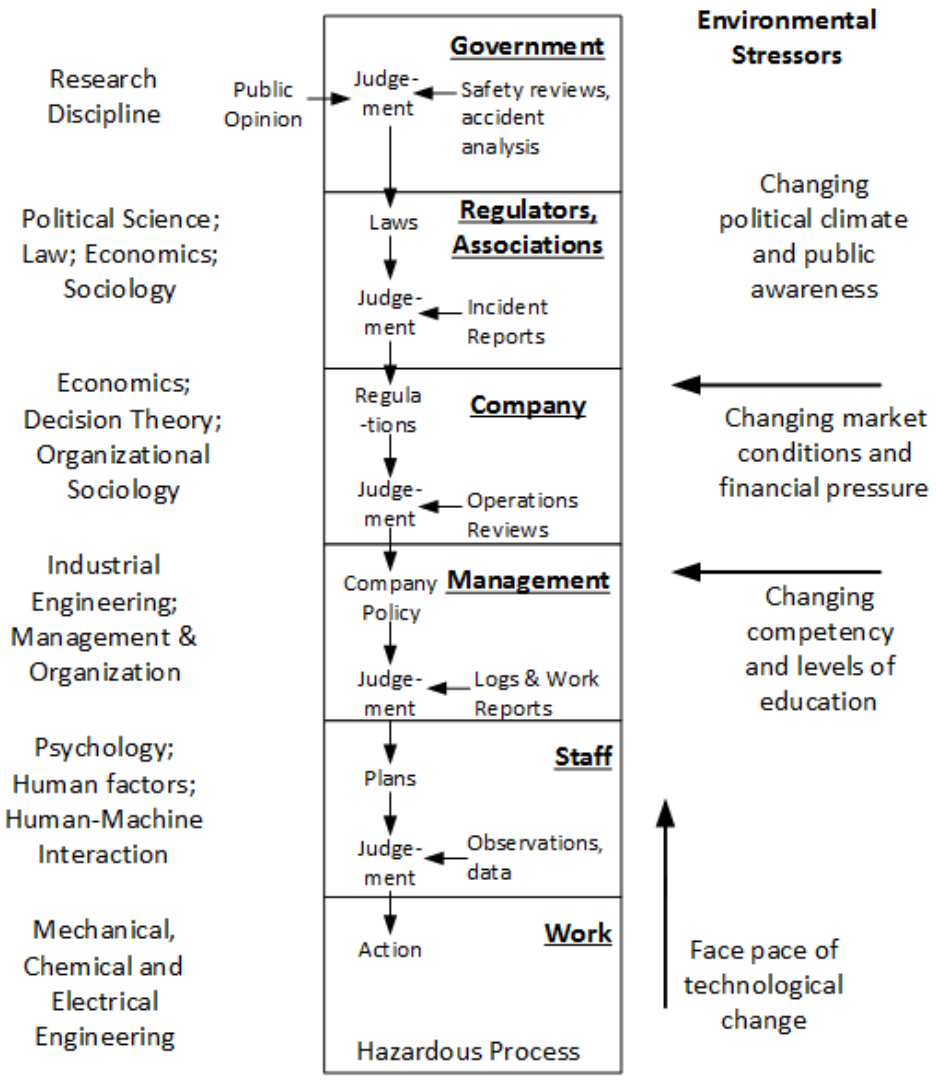


Figure 2-3 The hierarchical system model adapted from (Rasmussen, 1997)

The System Theoretic Accident Model and Process model, STAMP, introduced in (Leveson, 2004) by Leveson stresses the view of system safety as a control problem. STAMP includes details on the possible hazardous control actions, to which an accident can be traced.

The STAMP based Process Analysis (STPA) gives further hazard identification principles and guidelines in the safety control process. A basic assumption of STPA on accident

occurrence is that accidents originate from the erroneous or ineffective interactions along the loop of control, and preventions of accident require enforcement of control constraints at each stage of control (Leveson, 2004, Leveson 2011).

2.2 Safety Assurance in ATC

2.2.1 Improvement of infrastructure and separation standards

To control air traffic, ATC systems rely on Communication, Navigation and Surveillance (CNS) infrastructures to accurately track aircraft positions, guide flights along their planned routes, and collaborate with pilots, during the course of flights (Varon, 2000, Nolan, 2010). Detailed descriptions of CNS equipment and evolution are given in (Nolan, 2010). In the early days of ATC, uncertainties of aircraft positions using the primitive CNS systems are basis for the Reich model, which derived aircraft separation standards (Brooker, 2002a, Xu et al., 2008), and are still in use today. See Figure 2-4.

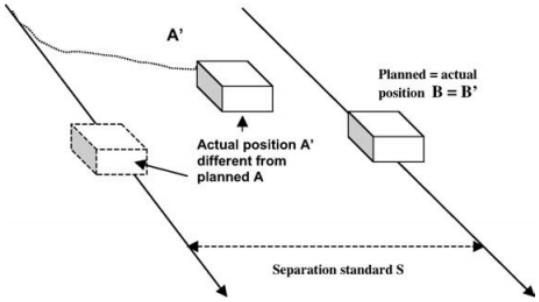


Figure 2-4 Reich Model (Brooker, 2002a)

ATC development tends to follow a bottom-up approach. Blom et al. pointed out that ATC/ATM improvement plans are organized around local features, e.g. improvements of the automation tools, the controllers/pilots and their human machine interfaces (HMIs), and the advanced procedures (Blom et al. 2001). Changes of local features will have an

impact on system's behavior, yet the enhancements of their particular functions do not guarantee system level safety performance improvements (Rasmussen, 1997).

Since the development of Reich models, surveillance environment has improved significantly. The separation standards derived from Reich model no longer stand on the same assumptions. For example, the guidance and navigation technologies have been greatly advanced since then, and the uncertainty of an aircraft's position has been significantly reduced. Figure 2-5 illustrates the change of contributions of aircraft position uncertainty to separation standards comparing ATC en route radar systems of 1950s and the modern day.

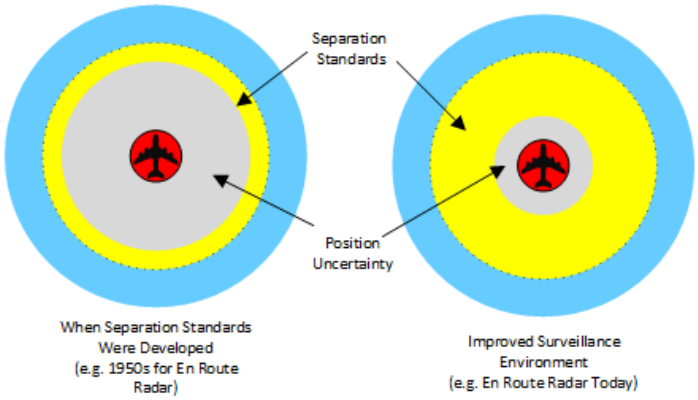


Figure 2-5 Aircraft Position Uncertainty Comparison (Hansman and Odoni, 2009)

The new safety challenges of modern complex engineering system like the ATC are summarized in (Leveson, 2011), including rapid development of new technologies, changing nature of accidents, and increased complexity and coupling. In ATC, safety is a result of interactions between a variety of system elements including human operators, procedures, and technical systems all of which are highly distributed. Safety should therefore be understood from the system level (Blom et al., 2006).

2.2.2 Redundancy and Defense in Depth

Redundancy is another common strategy for ATC safety assurance. Redundancy allows a reliable system to be built on “unreliable” components (Shrivastava et al., 2009). In ATC developments, new components are added to existing systems rather than replacing them. For example, primary radar was followed by secondary surveillance radar and variant “monopulse” radar (Brooker, 2008). This technique is also termed defense in depth, in that errors or occurrence of hazards will need to penetrate through layers of defenses to cause accidents.

Defense in depth originated from military tactics (Bass 2001). Instead of defeating an attack with one strong defense line, the defense in depth technique use several (weaker) layers of defense to delay or reduce the effect of attack, before an counter attack action can be taken. Such technique is also seen in other industries such as the nuclear power plants.

The use of redundancy and defense in depth result in a redundant, layered and complex system structure. Since ATC, like other safety critical systems, relies on the proper decisions of human operators/controllers, opaqueness due to system complexity will complicate the decision making process, and render system prone to human errors (Saleh and Bakolas, 2009).

Additionally, interactions between the different means of control are not fully known or tested, owing to the scale, complexity and lack of experimental environments in ATC safety research. Saleh and Bakolas proposed to compensate this shortcoming with the

concept of system controllability and observability borrowing concepts from the control theories and discrete event system theories (Saleh and Bakolas, 2009).

2.3 Safety Assessment in ATC

Safety assessments can be generalized into two categories, qualitative and quantitative. Qualitative safety assessment identifies the hazards and consequences of different failure modes, whereas quantitative safety assessments quantify the severity of the consequences and express the severity in terms of a predetermined safety metric such as risk.

2.3.1 Qualitative Safety Assessment

From studies of a variety of industrial safety issues, many hazard identification techniques are available both in practice and in research (Geisinger, 2003). Common approaches include Failure Mode and Effect Analysis (FMEA), Preliminary Hazard Analysis (PHA), Common Cause Analysis (CCA), and HAZard and OPerability study (HAZOP). Some of these techniques have been tested and adopted in the air traffic control system safety assessments (Dunjó et al., 2010a, Becker et al., 1997).

FMEA: Failure Model and Effect Analysis was one of the first systematic techniques for failure analysis and hazard identification. The method systematically postulates component failures and identifies the resultant effects on system operations, for as many components, assemblies and subsystems as possible. This process is based upon and recorded on specific FMEA worksheets. A successful FMEA needs to identify all significant failure modes for each contributing element or part in the system (Stamatis, 2003).

Although widely used for systemic qualitative system assessments, FMEA is generally considered a basic approach that cannot identify complex scenarios that involve multiple failures (Shebl et al., 2009, Franklin et al., 2012). The method also lack validity when used in isolation, owing to scoping and organizational boundaries (Potts et al., 2014).

HAZOP: A HAZard and OPerability study (HAZOP) is a structured and systematic examination, usually by a multi-disciplinary team, of an operation process to identify and evaluate hazards arising from deviations in the process. Deviations are marked by guidewords (MORE OF, LESS OF, NONE, REVERSE, PART OF, AS WELL AS, OTHER THAN…), to the process variables or parameters (flow, pressure, temperature, level, composition…) that are important to the process safety.

An example use of HAZOP is shown in Figure 2-6. The completeness of this hazard identification relies on team members' "intuition and good judgement" and the "climate of positive thinking and frank discussion" at the meetings (Site, 1994).

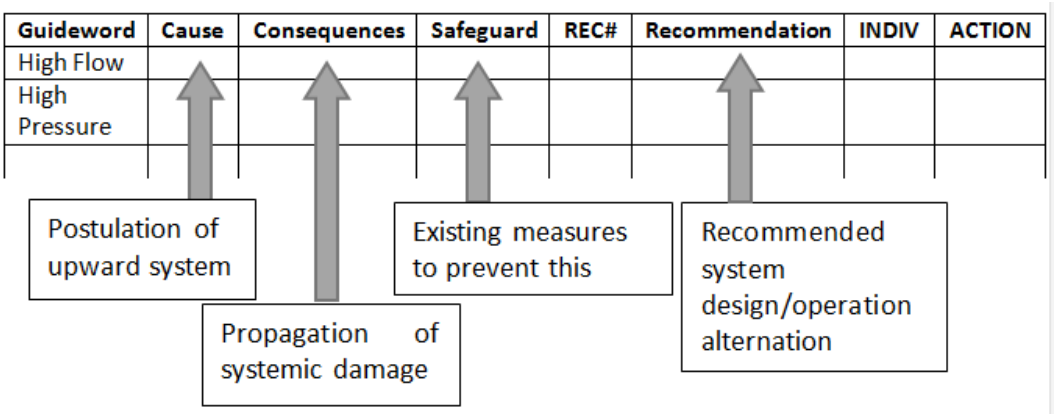


Figure 2-6: HAZOP worksheet example

An example of HAZOP use in ATC is given in (Leadbetter et al., 2001), which applied the method to identify human error sources including the user interface and cognitive process.

The HAZOP approach has a focus on operational and managerial hazards, rather than mere component failures. It however, requires multidisciplinary expert knowledge, and depends on identification of system mode and deviation modes (guidewords) (Dunjó et al., 2010b).

STPA: STAMP based Process Analysis (STPA) is a hazard identification method based on STAMP. From the operational control model in the STAMP hierarchy, STPA gives further guidelines on hazard identification in the "control" process (Leveson, 2004).

A basic assumption of STPA on accident occurrence is that accidents originate from the erroneous or ineffective actions along the loop of control, and prevention of accident requires enforcement of control constraints at each stage of control. The STPA generalize three groups of hazards: 1) inadequate enforcement of constraints, 2) inadequate execution of control action, and 3) in adequate or missing feedback (Leveson, 2004). For a complete list of control hazard types, see (Leveson, 2004) and (Leveson, 2011).

2.3.2 Quantitative Safety Assessment

Quantitative safety assessments are important to system designers and policy makers; it can be particularly valuable at early stages of a system's development to filter design concepts, on the safety ground (Blom et al., 2001). System safety needs to expressed and communicated with safety metrics.

2.3.2.1 Methods

A variety of quantitative safety assessment methods can be found in literature (Netjasov, 2010). Blom (2006) distinguishes static safety assessment techniques from dynamic ones. Static assessment methods are often combined with qualitative hazard identification methods and use of the hazard probabilities. Such methods include Fault Tree Analysis (FTA) (Andrews, 1998, Clemens, 2002, Ericson and Ll, 1999, Lee et al., 1985, Tanaka et al., 1983), Event Tree Analysis (ETA) (Andrews and Dunnett, 2000, Baraldi and Zio, 2008, Kenarangui, 1991), Bayesian Belief Network (BBN) (Gran and Helminen, 2001, Trucco et al., 2008). The dynamic techniques directly model the behavior and dynamics of a system. Examples of such include Petri Nets Analysis (PNA) (Liu and Chiou, 1997, Leveson and Stolzy, 1987), Multi-Agent Modeling(Blom et al., 2006), Markov chain modeling, and dynamic event trees (Durga Rao et al., 2009).

Since the focus of this research is on the safety metrics, rather than the methods, the details of these methods will not be discussed. See (Netjasov, 2010) for a review of risk modeling and safety assessment methods in aviation.

2.3.2.2 Safety Metrics

Frequency based metrics: Frequency based metrics are used to describe the past accident data and project future trends of system safety performance. The most widely used safety metric is “the count of accidents (fatalities, loss of aircraft) per unit time/unit distance”. Other similar metrics include number of incidents per million flights, number of fatal accidents or number of casualties per year, mile flown, or passenger

hour (Brooker, 2004, Netjasov, 2010). The current level of safety in the current ATC is about 1 disastrous accident per 10^7 flight hours (Savage, 2013). The denominator has a dramatically long period because air traffic accidents are extremely rare. In the US, over the span of 10 years (2005-2014), there was a total of 12 fatal commercial aviation accidents, and world wide, 72 (Airplanes, 2015).

Probability based metrics: Probabilistic metrics focus on specific scenarios and probabilistic risk of accidents. Many simulation based safety assessment approaches, such as petri nets (Netjasov et al., 2013), Monte Carlo Method (Baraldi and Zio, 2008, Stroeve et al., 2009), mutli-agent models (Stroeve et al., 2013, Blom et al., 2006) use "risk" as the safety measure. In these cases, the risk is synonymous to the "probability of occurrence" of the concerned accident. Generally, risk estimate will involve both likelihood/probabilities and severity associated with accidents (Siu, 1994). Variations of expertise in the analysts directly affect the completeness and accuracy of results.

Another critical challenges to implement the probabilistic metric, is that safety critical systems are almost perfectly safe (Amalberti, 2001). As mentioned previously, the target values of probabilistic risk are at the order of f 10^{-7} to 10^{-10} per aircraft flight hour (Blom et al.). A single accident will introduce significant fluctuation to the safety levels using such metrics (Brooker, 2002a).

Comparison in the quantitative safety assessments of ATC is between two extremely small numbers. When both numbers have an uncertainty of more than one magnitude of the estimated values, say between $10^{-10} \mp 10^2$ and $10^{-12} \mp 10^2$, the comparison will be questionable and unconvincing.

Metrics of incidents: Some metrics use incident data to infer safety. Close proximity indicator (CPI) for example measures the closest distance of two aircraft in the case of loss of separation. Similarly, Severity Scores, Actual Separation Breach, Incident not resolved by ATC (INRA) are other metrics that use incident data (Brooker, 2007).

The drawback to incident data is that incident data are incomplete due to various complication factors, e.g. pilots fear of penalties. In 2011, Euro Control Safety Regulation Commission estimates that only half of incidents are reported.

2.3.2.3 Safety Performance and "System Control Capacity"

In Rasmussen's view of safety as a control problem (Rasmussen, 1997), he first brought about the concept of "loss of control boundary". Rasmussen promoted that safety management should make the loss-of-control boundary explicit and visible to the actors, and "increase the margin" from this boundary. The margin arguably is a manifest of a system's control capacity.

The categorical counterpart of control capacity, "controllability", that whether a system has some control capacity or not, was first introduced to the safety engineering context by Saleh and Bakolas (Saleh and Bakolas, 2009). Controllability has its origin from control theory (Kalman, 1959, Klamka, 2013). This concept of controllability and observability was brought into the safety engineering context to compensate the effect from using "defense in depth" safety strategy in safety critical systems. Controllability is defined as the ability of bringing an accident initiating event back to the "safe zone" in the system's state space. Observability is defined to be a system's diagnosis ability of

"pathogens". Controllability and observability are used as principles to examine the processes, rather than safety measures.

The use of controllability as a safety measures is also seen in the EuroControl SAfety Regulatory Requirement (ESARR) Advisory Material 2/ Guidance Document 5 (EAM 2/GUI 5), as shown in Figure 2-7. They view controllability (highlighted in Figure 2-7) as the system capacity available to resolve an impending accident. However, it was not indicated how this controllability should be evaluated.

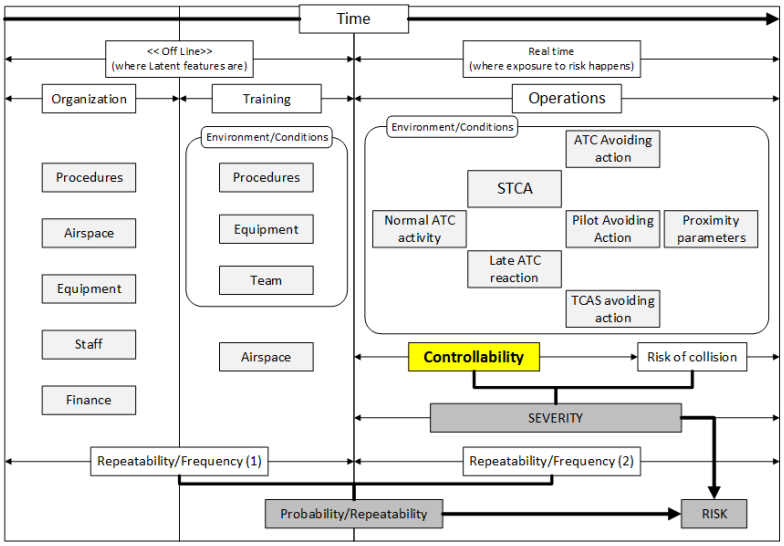


Figure 2-7: Air Traffic Control System Risk Analysis (SRC, 2005)

Despite the wide acceptance of correlation between system control and safety, and implications of a 'safety margin', by the discussed research, system control capacity, as a quantifiable system property in the safety engineering context, has not been formally defined; and the relation between system control capacity and safety performance has not been formally established, to the best of my knowledge. The research is put forth to

address these gaps, and examine the viability of using system controllability as a safety performance metric.

2.4 Summary

This chapter reviewed the current research on ATC safety and safety assessments. The research on system control and system safety, as well as system control capacity and safety performance is placed in the ATC safety assessment and management research context. Two observations are made from reviewing the existing research. First, there is no formal definition of a quantifiable control effectiveness measure, which we call "control capacity" in the context of safety engineering. Second, although the correlation between system control and safety have been well studied and documented in the literature, this system control capacity as a potential system safety performance measure was never discussed. The literature review reinforces needs to explore the research questions proposed in Section 1.3. The following chapters detail the investigation conducted in search of answers to these research questions.

CHAPTER 3. DEFINITIONS AND THEORIES

This chapter builds the theoretic basis for assessing system safety performance with system control capacity, including formal definitions and assumption statements of this research. Sections 3.1 to 3.3 detail peripheral ideas, concepts, and theories, which are important in the understanding of system control capacity. Control is abstracted as the interactions between a system, the *controller*, and a process, the *controlled*. Variations on both sides are demonstrated to have an impact on the system's control capacity. To understand the controller, the control system is placed first in the general system theories for the most fundamental features, then a discussion on complicating factors that affect its control capacities. A process is generalized as a state transition from the current state to a desired state. Founded upon the proceeding sections and from the literature, formal definitions of system control capacity and the research assumption statements are given in Sections 3.4 and 3.5.

3.1 The Concept of Control Capacity

3.1.1 Variation of control capacity

The Oxford dictionary defines control as "the power to influence or direct people's behavior or the course of events". In engineering systems, the verb "control", refers to the exercise of such power, which is achieved using a control system.

Control systems take many forms, but ultimately can be abstracted as a controller, a controlled process and the bidirectional interaction between them. As is illustrated in Figure 3-1, the controller acts on the process to change its characteristics, e.g. pressure, speed, or power output; the controlled process exhibits information regarding the effect of the controller's acts, which is necessary for serving the purpose desired by the controller, as it guides the control to move towards the desired outcome.

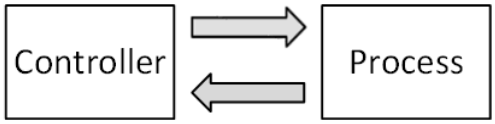


Figure 3-1 Control in simplest form

Capacity is the ability or power to do, experience, or understand something (Oxford Dictionary, 1989). Combined with the definition of control, the term "control capacity" therefore is the power of a "control system" to influence or direct people's behavior or the course of events.

In the safety engineering context, I define the set of "active" controls as those control activities that are driven by actively monitoring and influencing the controlled processes. In contrast, the set of "passive" controls does not change the course of events; they simply contain and control the severity of accident consequences. For example, in a nuclear power plant, active control is through the operator's constant monitoring of reactor outputs, whereas an example of passive control measure is the concrete walls that shield the reactors.

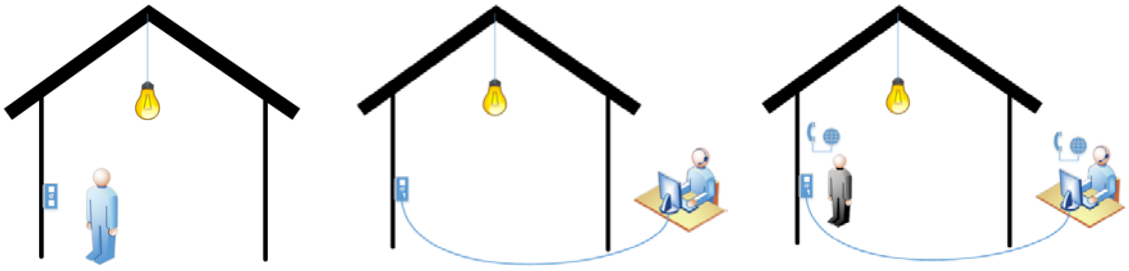
In principle, for an active control means to have an effect, there should always be some power of the system to influence the development of the process, and hence some control capacity. Since some systems are safer than others, there should also be variations of systems' control power/control capacity in attaining their control objectives. In other words, some system should have higher control capacity than others.

3.1.2 Contributing factors to control capacity variations

Consider Figure 3-2, where three systems are used to switch a light on/off in a room. System I has the controller standing next to the light switch and the light can be directly controlled by flipping the switch. In System II, the controller is away from the room and can access the switch through instrumentation, e.g., using Labview. In System III, the controller can also control the switch via instrumentation; additionally, the controller can command an operator in the room to switch the light through telecommunication.

All three systems have some control capacity over the process: switch Light from ON/OFF. The following observations can be made however to distinguish control capacities among the different systems, over this same process.

- System I has more direct access to the switch and hence the light.
- System II requires more intermediate steps between the controller and the switch, compared to System I; and compared to System II it has not alternative control means, should the instrumentation fail.
- System III, compared to System I, does not have as much direct access to the switch, but has more than one *means of control*.



a) System I b) System II c) System III

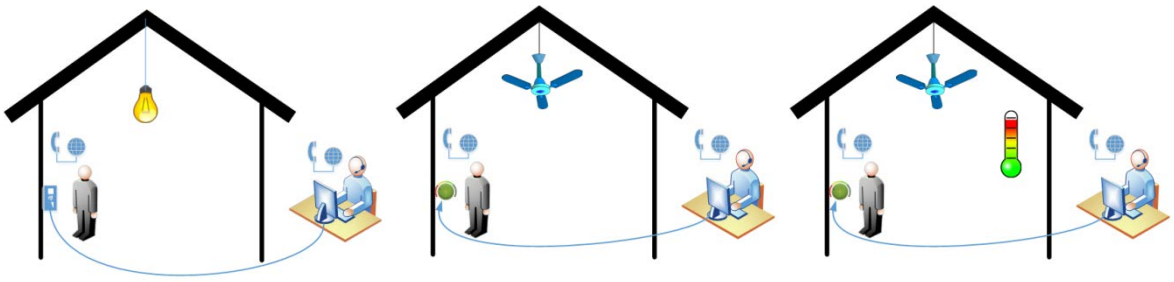
Figure 3-2 Light Switch Systems with Differing Controllability and Observability

In comparing the three control systems intuitively, over the same process, System II appears to possess the lease control capacity, since it is not as direct as System I, and does not have as many control means as System III. However, between Systems I and III, it is unclear which one is more capable of control. The system property that differentiates the three control systems in their control activities, control capacity, is the subject of this research. If quantification of control capacity is possible, the comparison between the different systems can then be carried out with confidence.

The above thought experiment also uncovers two factors which play important roles in determining the control capacity of a system:

1. Directness: Control capacity decreases as the control "route" becomes less direct.
2. Redundancy: Control capacity increases as the number of control "routes" increases.

Now if the same system is used to control different processes, how do the control abilities vary? In answering this question, compare three control systems illustrated in Figure 3-3.



a) Process I b) Process II c) Process III

Figure 3-3 Controllability varying with processes

In Figure 3-3, the same control mechanism is applied to controls of three different processes. Process I is the control of a Light on/off as is in Figure 3-2. Process II is the control of the rotational speed of a fan. Process III is the control of a room's temperature, by increasing/decreasing air circulation in the room.

Compare the three processes. Process I appears to be easiest to control, since there are only two possible states, i.e. {LIGHT ON} and {LIGHT OFF} and the control action is apparent.

Process II, compared to Process I, is a more complex task. To change fan speed from one to another will require more delicate operations of the control dial than that on a switch.

Process III is the most complicated one to control. In this process, states are measured in temperatures. To change room temperature from one value to another, the fan is used to increase airflow. This process is even more demanding in the operation task of dials.

All three control systems have a second means of control via telecommunication. In this case natural language will have to be used in delivering control commands. In Natural languages are subject to misinterpretations, and more prone to errors than coded

commands (e.g., 1 = ON, 0 = OFF). The more complex the control task, the more demanding it is in avoiding miscommunication.

In summary, this example demonstrates that for a system to control over different processes there is also distinguishable variation in the control system's control capacity. The factors that affect control capacity, uncovered by this simple example include:

1) Size of the state space, e.g. Process I compared to Process II;
2) Size of solution space, e.g. System + Process I, compared to System + Process II;
3) Complexity of the control command messages e.g. controls via instrumentation compared to controls via telecommunication.

Thus far, I have discussed and demonstrated the general concept of control capacity and its variations due to system and process differences. The concept is generalizable to any control systems with different control purposes.

3.2 System

The term "system" is used in nearly all scientific and engineering disciplines, and refers to artifacts of various forms, constituents and scales. General Systems Theory (GST) is a field that unifies studies of systems across disciplines and works with issues common to all systems.

As per GST, a general system is any construct or collection of elements whose interactions fulfill certain objective(s) or purpose(s) (Hall and Fagen, 1956). While controls can be generalized to any system including biological and ecological, here the focus is on systems designed and engineered by humans to serve humans. These systems are made of hardware, software, equipment, facilities, personnel, processes, procedures and so forth, all of which are referred to as system elements. A system

element here is a generalized term for all possible system constituents, regardless of its level of abstraction or form of existence.

Here, the term “system” is used both in the control system that consists of the controller and the controlled process, and the controller by itself. The limiting word “control” before the system in the former is purposely used to differentiate between the two. The controller itself is referred to as a system to be consistent with the subject of this study, the ATC. As per FAA definitions, the ATC “system”, provides a service to ATC users, air traffic in its controlled space. The control system to be analyzed will include both the ATC system, and its controlled traffic which in this research is referred to as the “process”.

Control capacity varies as the system or controller change. To understand the dimensions of systems’ variations, we resort to GST for the basic traits of a system.

3.2.1 Basic views of a general system

In GST, there are three fundamental views of any given system: structure, function, and attributes (Bertalanffy, 1968, Torokhti, 1975, Skyttner 2005 and Boulding, 1956).

Structure View: In the structure view, a system is described as a collection of objects and relationships. Symbolically (Bertalanffy, 1968, Klir 1991, and Skyttner, 2005),

$$S^s := O \times R \qquad (3.1)$$

Where *O* is the well-defined object set of the system, that is, its elements are enumerable.

R is the well-defined relationship set of the system;

"×" defines the Cartesian product between two sets. The Cartesian product of two sets A_1 and A_2 is defined as:

$$A_1 \times A_2 = \{(a_1, a_2): a_i \in A_i, \forall i\} \qquad (3.2)$$

Visually, this definition system can be interpreted as a graph, shown in Figure 3-4.

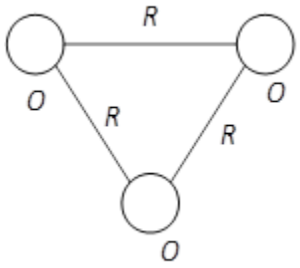


Figure 3-4 Structural view of a system

Function View: In some cases, it is useful to pack part or all of the system as a function that expresses input-output behavior at a higher level. The input-output relationship is analogous to a black box, whose internal mechanism is not, or need not to be known by its observers or users. The input and output patterns will serve certain useful and consistent purposes, and thus a function.

Symbolically, the function view describes a system as (Torokhti and Howlett, 1975):

$$S^F: I \longrightarrow Y \qquad (3.3)$$

Where, I is the input set and Y is the abstract output set.

The functional view of a system is illustrated in Figure 3-5.

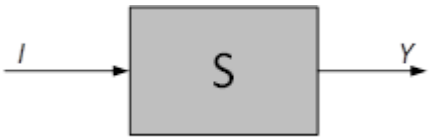


Figure 3-5 Functional view of a system

Attribute View: In the attribute view, a system can be abstracted into a set of *attributes* that characterize the system. System attributes can be logical such as the state of a light: "ON" or "OFF", or quantifiable, such as the speed of a train. Some attributes can be directly measured e.g. temperature and pressure, while others are indirect manifest of system characteristics e.g. reliability and stability.

Symbolically, a system from the attribute view can be expressed as (Kalman, 1959) and (Torokhti, 1975)

$$S^A \coloneqq \{x_1(t), x_2(t), \dots\ x_n(t)\}. \tag{3.4}$$

Where, $x_i(t)$ is the value of attribute x_i at time t.

The three basic views of a general system are considered enriched and extended definitions of a system, following its general definition in Section 3.2. Concepts elaborated and illustrated here including function, attributes and structure will be used in the proposed metrics and method of this research. Additionally, they provided basis for understanding of the nature of variations of different systems.

3.2.2 Complexity of ATC systems

An ATC system will vary as its structures, functions, and attributes change. Controls of ATC also rely on other structural, functional and attributes-related complex features that affect the system's control capacity.

Stroeve pointed out factors that makes an ATC system complex including: 1) The number and types of entities (human roles, technical systems), 2) number and type of interdependencies between entities, 3) degree of geographical distribution, 4) type of dynamic performance for components, and 5) number and types of hazards, i.e.

situations/conditions that may decrease the level of safety (Stroeve et al., 2009). The five features may be traced back to the basic views of a system, but they are discussed separately here to illustrate the complexity of the problem.

Hierarchy: The ATC system has a distinguishable hierarchy in its organization and operations directing air traffic (Haraldsdottir et al., 2001), from national flow planning, to center/facility planning and sector control. Hierarchy is the organization of system elements, which ranks some above others, e.g. according to levels of authority. The STAMP model illustrated a hierarchical pattern of hierarchy in social-technical systems where higher levels, e.g. regulatory agencies have higher authorities over lower levels, e.g. management. The hierarchy is a system structural characteristic, but bears other distinctive and complex features as well, e.g. levels of abstraction (DeLaurentis, 2005).

Integration of humans and automations: In ATC, humans and automations are closely knit to share responsibility of surveillance and control of the system. Automated systems have high computing speed, and can store and process large amounts of information (Wickens, 1998). In comparison, humans have the advantage of being flexible, tactical and more adaptive to dynamic and complex processes.

The integration of human and automation is a highlighted feature of the group of social-technical systems; it brings many challenges to the operations and understanding of the such systems, including introduction of unknown safety hazards (Wickens, 1998). In the Uberlingen mid-air collision accident for example, the conflicting commands from the automated collision avoidance system and the ground controller led to the

uncoordinated collision maneuvers and consequently, the collision accident (German Federal Bureau of Aircraft Accidents Investigation, 2002).

Interdisciplinary: Control in ATC is the integrated effort of all system elements including hardware, software, personnel, and facilities. The human components, in particular, play important roles in assuring the ATC safety. The comprehensive understanding of the system elements' and the overarching system's behaviors requires a cross-disciplinary effort from a spectrum of studies on system theories, psychology, and economics, to name a few (Rasmussen, 1997).

Geographically distributed: The ATC system infrastructure is highly distributed. Like many other systems, the decisions made remotely need to be delivered to the operators or actuation units. Likewise, the information at the end of operations needs to report to the decision makers often at a different location. Missed, distorted, late or ambiguous information can lead to failed functions and even cause accidents to occur (Leveson, 2004). Therefore, communication is particularly important to accurately and timely facilitate the interaction between system elements.

3.3 Process

Control objectives: A Controller by itself does not have safety concerns; whether a control system is safe or not depends on the processes to be controlled. In safety control in particular, the act of control is activated by the need of the system at the start of the control, the initial system state. If the state is projected by the controller as "unsafe" or hazardous, the controller must then move the system back to a safe state.

Safety control, in general is to accomplish this transition from a perceivably unsafe state to a safe state.

This type of "event-driven, continuously having to react to external and internal stimuli" system can be viewed as a reactive system (Harel, 1987). The ATC system can be classified as a reactive system: "internal stimuli" occurs at critical sections of a flight, e.g. taking off, landing, and crossing borders of a controlled sector; external stimuli are beyond routines, although still expected, e.g. resolve conflict between two aircraft, and assist pilots in distress.

States: A state is the particular condition that someone or something is in at a specific time (Davis, 2005). The selections of the conditions and criteria for determining the states based on the conditions are arbitrary. It may vary with the viewers' perceptions, and purposes of the state selection.

In control theories, system states are expressed by a set of state variables(Hall and Fagen, 1956). These variables can be continuous, discrete or descriptive, but in general, they should characterize behaviors of the control system, or the processes to be controlled. The control theories define the each unique combination of the state variables to be a system state. This is equivalent to defining state for a time instance.

For safety controls, as long as the destination state is considered safe, defining value differentiated states may not be necessary. In this case, the state is characterized by a zone of state variable values. For example, to resolve a conflict between two aircraft, there may be many solutions: one aircraft climbs, the other descends, or the other way around; they can both make a turn to the opposing directions on the same flight level,

etc. The resultant state does not require the aircraft to be at certain positions. To express states as zones, both the choices of the state variables and the values of the variables need to attend to the purpose of the analysis.

The set of all possible states constitutes the state space. Bahill pointed out that a system state has two basic properties 1) it characterizes the system for a period of time (representiveness) and 2) it is different from other states (Bahill, 2011). Additionally, at any given time, the system is in exactly one of the states from the state space. See (Wasson, 2011) for a survey of the definitions of system states.

These three characteristics of system states/state space representiveness, mutual exclusiveness and exhaustiveness are necessary criteria for valid definition of system state space.

Dynamics and Control: As previously discussed, a controlled process can be abstracted and described as a state transition, from the current state to a target state, e.g. to move from an unsafe to a safe state. The driving 'forces' of transitions, in the general sense, are necessary inputs/events and system conditions (Ptolemaeus, 2014). The study of state transition trajectories, in relation to the influence, or driving forces, is the subject of system dynamics. System dynamics can be described with equations, indexing tables, linguistic descriptions or graphs.

When the transition or trajectory of system states can be deliberately manipulated to yield desired utilities or functions, the system is controlled. In control systems, the decision maker subjectively aligns the state transition with the purpose and function of the physical process, through intervention to the system dynamics.

3.4 Define system control capacity in the context of safety engineering

Control capacity is a general property to all control systems. From the safety perspective, we define control capacity as follows:

Definition: Control capacity is the extent for a control system to withstand performance deviation in acquiring its control objectives.

This definition underlines the assumption that if performances of the system or parts of the system do not deviate, there would be no safety events. System elements are designed to serve purposes and required to meet certain constraints, without which the system will not function as intended. Disturbance, disruption, degradation and faults drive the system elements to deviate from the designed performance, which creates the conditions for an accident to occur. The definition is applicable to any general system, and not limited to ones with mathematical models.

Control capacity with this definition is an indicator of system's tolerance of performance deviations. Referring to the dictionary definition of control, control capacity measures the "power of influence", by examining its tolerance of losing this power.

3.5 Relate system control capacity to system safety performance

Based on research on system safety and control, we elicit three statements that constitute the theoretic basis for using control capacity to measure a system's safety performance.

Assumption 1: *A safe system must always possess some control capacity (Saleh and Bakolas, 2009).*

A safe system is first controllable. Conversely, an uncontrollable system is unsafe. This assumption was elicited from (Saleh and Bakolas, 2009) and also considered self-evident.

Assumption 2: *Control capacity is an indicator of the system's performance in the control of safety critical processes.*

To interpret this assumption, first define safety critical process as follows:

Definition: Safety critical process is state transition $p: X_I \rightarrow X_D$*, whose failure will result in unsafe consequence, where* X_I *is an arbitrary initial state, presumably undesirable and* X_D *a desirable state.*

Control capacity is a generic property common to any control systems, but when it comes to the control of safety critical processes, it becomes safety critical to stay controllable in despite of performance deviations. the system's control capacity, which measures the resilient power to performance deviation, is then an indicator of system safety performance. In other words, the knot that ties the quantitative measure control capacity and system's safety performance is the system's safety critical processes.

Assumption 3: *Between two similar control systems (of the same safety critical process), the higher the system's control capacity, the safer the system (Rasmussen, 1997).*

A safe control system should have some tolerance over performance deviation, and therefore some control capacity. Control capacity with our definition is an indicator of system's tolerance under performance deviation. For example, how much component failure can the system tolerate before loss of control, or how much delay can the system tolerate before loss of control?

The only thing that does not change is change itself. It is unlikely a system will act exactly as expected, due to changing environment, operational conditions, equipment degradation etc. There is always some performance deviation occurring throughout the life cycle of a system. Safer systems should be more resilient to fault, disturbance, degradation and disruption, than less safe systems. Control capacity is defined as a measure of such ability. Therefore, in the control of safety critical processes, higher control capacity is always desired for safety, if allowed by other system constraints.

These three assumption statements constitute the theoretic basis for justifying the use of control capacity as a safety performance measure, and following research to develop metrics and methods of control capacity based safety assessment of air traffic control systems.

3.6 Summary

In this chapter, the concept of system control capacity is formally defined and decomposed. System control is divided into two parts: the system that controls and the process to be controlled. When the controlled process is safety critical, system control capacity is relevant to system's safety performances. Comparing systems with different configurations requires quantification of system control capacity. The following chapters describe metrics and methods for evaluating system control capacity, for quantitative safety assessments in ATC.

CHAPTER 4. ASSESS SYSTEM SAFETY WITH CONTROL CAPACITY-PART I: METRICS

Recall Figure 3-1, where control is abstracted as the interaction between the controller and the controlled process. Also in Chapter 3, arguments are made that variation on either side will result in changes of a system's control capacity, "the extent to withstand performance deviation".

On the controller side, the two most common and concerning performance deviations are failure and faults. Failures are losses of function, and faults are malfunctions. This aspect of control capacity is selected to represent the control mechanism quality, and measured by a Probabilistic System Control capacity (PSC) metric. When the controlled process is concerned, timing is an important and holistic aspect for safety control (Leveson, 2004, Tian et al., 2012, Landry, 2012). The performance deviation related to timing is delay. For this reason, a Temporal System Control capacity (TSC) is introduced here to represent the variation of processes as well as impacts of such variations on systems' control capacities

.

4.1 Probabilistic System Control Capacity

4.1.1 Definition

Definition: A system K's Probabilistic System Control capacity (PSC) over a process $p: X_I \rightarrow X_D$, $PSC(K, p)$ *is the probability of System K to preserve at least one means of control, in the presence of partial failures or faults:*

$$PSC(K, p) = P(\cup F_i) \qquad (4.1)$$

Where,

X_I: Current state of the process;

X_D: Desired state of the process;

F_i: The event "to retain Means of Control i".

Since the probability of preserving at least one means of control is equivalent to one minus the probability of losing all means of control, which is sometimes easier to obtain, we can use either formula to evaluate $PSC(K, p)$:

$$PSC(K, p) = P(\cup F_i) = 1 - P(\cap \overline{F_i}) \qquad (4.2)$$

Where $\cap \overline{F_i}$ denotes the event "all means of control are lost".

Process $\boldsymbol{p}$: $\boldsymbol{X_I \rightarrow X_D}$: The concepts of state, state transitions and safety critical processes are detailed in Chapter 3. To reiterate, system states are arbitrary characteristic descriptions of system at different times or under different conditions. The state space of a system is the collection of all possible states. Within the system's state space, the system can move from one state to another. In this respect, processes are state transitions. Safety critical processes are a subset of all possible state transitions, whose failures will result in unsafe consequences.

Means of control: A means of control from a function perspective is the minimal set of functions capable of controlling the process, and from a system structure perspective, the minimal set of system elements.

This definition is more suited for systems with complex architectures, which provide more than one means of control. As discussed in Chapter 2, redundancy is a commonly used strategy for safety assurance in safety critical systems, like the ATC systems.

Means of controls may or may not be independent from each other. In the design phases, a system will go through cost benefit tradeoffs. For costly changes, like ATC infrastructures, it is but reasonable to design redundancy for only "critical parts" of the system whose performance increase is most necessary and rewarding. "Non-critical" parts of the system will be shared by the different means of control. This cost-effective safety strategy determines that the resultant system architecture will have coupled control means.

A second reason for multiple means of control is the need for collaboration in the overall controller-process task. The US ATC, at the highest level, is responsible for air traffic in the national airspace. The hierarchy of the control structure divides the overall task to manageable sizes for individual controllers. In this task division, multiple controllers may be assigned to control the same process; this is referred to as overlapping control (Leveson, 2004). The incentive for overlapped responsibility assignment is to allow cross-examinations, which reinforces the safety guards. Facilitating the collaborations between the overlapped responsible parties may require overhead management and additional procedures.

Additionally, infrastructure systems like ATC evolve over time. The addition of new components, procedures and concepts of operations are not in some original design plans. As discussed in Chapter 1 and 2, there is a tradition to overlapping new components to the current system to create a “defense in depth”, against accidents. For these reasons, redundant control means are intrinsic to the design and development of safety critical systems.

Failure and faults: We distinguish loss of control from effectiveness of control in that the former concerns loss of the function, and is often the result of failures, while the latter may include malfunction as well, often resulting from faulty interactions between system components.

The causes of performance deviation cannot be exhaustively enumerated when all system types are considered, but a general taxonomy of faults can be theorized. With the system safety be considered as a control problem, Leveson summarized the factors contributing to ineffective control under the STAMP model assumptions including 1) inadequate enforcement of constraints, 2) inadequate execution of control actions and 3) inadequate or missing feedback (Leveson, 2004). The more detailed list of causes to control inadequacies is found in (Leveson, 2004).

Probabilistic Risk and PSC: Risk is often expressed as the probability of the occurrence of a specified undesirable events (failure, accidents etc.) and estimates of their consequences (amount of losses, damages, injury, etc.). Probabilistic risk refers to only the probability or likelihood of the consequence.

The measure PSC and probabilistic risk share many similarities. If evaluated for the same undesirable consequence such as a collision accident, PSC can be an estimate of collision risk. The difference however is that PSC examines from the control perspective; it builds on the assumption that system safety and accident prevention is a control problem. In contrast, evaluation of probabilistic risk does not specify how the "undesirable consequence" is prevented. For example, probabilistic risk can be derived from direct simulation of a system's behavior (Blom et al., 2006).

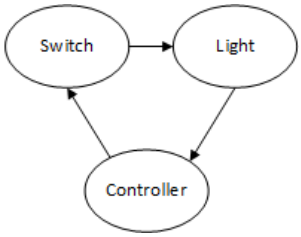
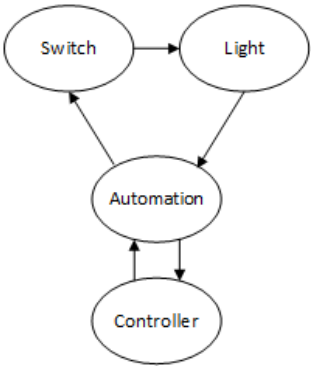
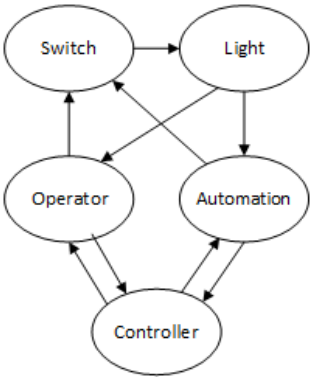
Estimated PSC: One challenge to quantitatively assess safety in safety critical systems is the lack of accident data. The proposed measure, PSC, as defined, requires knowledge of the probabilities of unlikely events, which are difficult to acquire. The values of the needed probabilities, in practice should be elicited from experts, resources, as are the case with other Probabilistic Risk Assessments methods.

The focus of this measure however is at the system level and interactions between different means of control. Therefore, the hazardous events are identified at higher levels as well, e.g. subsystem failure, compared to those needed in a PRA study, e.g. device failure. The values of probabilities for subsystem level hazard events are expected to be based on judgments. Since the primary utility of PSC is to differentiate the control capacities of different control systems, and not to obtain the absolute risks, the accuracy of these probabilities is therefore assumed secondary. This assumption will be tested in the case studies. For details, see Chapter 6.

4.1.2 PSC illustrated

This section uses an example to illustrates the concept and estimation of PSC. We apply the PSC definition to the example used in Section 3.1: controlling of a light from ON to OFF. From the illustration in Figure 3-1, first extract the control mechanisms on the process {LIGHT ON} → {LIGHT OFF}. Use the structure view of a system, this control mechanism is represented by directed graphs (Henley and Kumamoto, 1985), as shown in Table 4-1.

Table 4-1 Control Models of Simple Systems

System I	System II	System III
		

Since PSC considers the system's tolerance to failure and faults, naturally, the next step is to find possible faults and failures that may hinder the control process. A preliminary hazard identification (for fault and failure) provides that System I may have the following possible faults and failures: 1) wires between switch and light are disconnected. 2) Switch fails due to the physical mechanism and 3) the controller fails to

turn the switch. Similar “hazard identification” processes can be made to the other two systems.

From here, we may use PSC as the measure to assess, compare and identify the systems that are more resistant to failure. In System I and II, there is no redundancy, that is, if any of the system element fails, there will be no substitute counterparts. Whereas in System III, should the automation fail, the system may still manage to control the light via the operator. On the other hand, for System III, it is possible that combining Means 1: the error prone communications between operator and controller, and Means 2, the unreliable automation system, it is not as effective and dependable as the only control means in System I, having the controller directly control the switch.

To quantify the difference and make the comparison, the next step is to compile all identified hazards into a hazard list, rank these hazards according to their likelihood, and assign an estimated probability of occurrence to each hazard. Occurrence of a single hazard may or may not induce loss of control, the final step is to enumerate all possible combinations of hazards that does and find the likelihood for each of the system in losing all means of control, or preserving at least one means of control, or PSC.

The choice of probability values presumably will have an impact on the result. The most credible way to estimate these values is to use reliabilities of system elements if available, and resort to experts otherwise. Alternatively, impacts on the choices of the hazards can be proactively tested to find boundaries of the values when qualitative differences are observed in the comparison, and then determine which groups of values

are more aligned with empirical or historical data. This alternative approach will also give a level of confidence on the conclusion drawn from the comparison.

The example demonstrates the three basic needs of quantitative safety assessments: a control model, a list of hazards, and probabilities of hazards. Note in this example, the process to be controlled, {Light ON} to {Light OFF}, is obvious; however for more complex scenarios, a systemic approach to safety critical process identification may be desired. The previous chapter establishes the argument that relevance of control capacity to a system's safety performance depends on the "safety critical processes". Chapter 5 will elaborate on principles and procedures recommended to identify "safety critical processes", for quantitative safety assessments.

4.2 Temporal System Control Capacity

Delay is a common type of performance deviation in human-centered, and geographically distributed systems; in the control of "safety critical process", however, delays can result in unsafe consequences. A safe system should always tolerate some delays, this higher this tolerance, over all of its safety critical processes, the safer the system. A temporal system control capacity (TSC) measures a system's control capacity, when the performance deviation, delays are concerned.

4.2.1 Definition

Definition: The Temporal System Control Capacity $TSC(K, p)$ for a system K and over a process p: $X_I \rightarrow X_D$, $TSK(K, p)$, is the difference between the time available, t_A, and the time required, t_R:

$$TSC(K,\ p) = t_A - t_R \qquad (4.3)$$

Where, t_A and t_R are defined as follows:

Definition: time available t_A is the time difference between the time instance when the forming of an accident is detected by the controller and that when the accident occurs, should no control be applied.

Without control, the system state evolves over time following relevant physical, organizational, and regulatory principles. The time for the process to evolve from the current unsafe and undesirable state to a new safe and desirable state therefore can be calculated applying the principles.

Time available is calculated as a reference. In reality, when an unsafe process is detected by the controller, should a control resolution exist given the situation, a controller presumably will always attempt to resolve the situation. Therefore, the case that an imminent accident is detected and no control will be applied is extremely unlikely.

Definition: Time required t_R is the expected time consumption for a safety control to initiate and take effect in response to detection of a safety threat.

The core component in a control system is the controller. Safety control is often motivated by external stimuli or the episodic needs to handle a safety threat; it is initiated and deactivated by the controller. Under the no performance deviation assumption, time required is the nominal and expected time for a control action to be generated and to take effect. Each component involved in the control process will contribute to a portion of time required.

Traffic surveillance systems such as the radars collect positions of aircraft continuously. The control loop used for safety assurance will not be activated however until the controller recognizes any threat to safety. For example, the precursor event to accident that two aircraft enter collision course, under the control of air traffic controller, is often detected at the violation of separation standards. Before the relative distance of the two aircraft reach the standard separation, likely no control action will be initiated or taken.

Implications of TSC: In any case, time required t_R for a control action to take effect should be at least as large as the time available t_A, to assure safety. Preferably, the bigger the difference (TSC), the more capable the system is to handle delays during operations. Delay from the expected values at any part of the system is considered a performance deviation. In the control of safety critical processes, excessive delays are not acceptable. If the control fails to take effect in time, an unsafe event, or accident, will result.

From the design perspective, the system should be designed with positive TSC. This can be accomplished in a number of ways. First time required can be reduced through technological advancements such as upgraded infrastructure and inclusion of automation assistance. Another effective way to increase TSC is through early detection. This however is a conflicting goal to other constraints in the traffic control. For example, the increase of separation standards will allow early detection, but it will also reduce the airspace capacity.

With the measure of TSC, strategies should also be sought as to how to avoid excessive delays that may exhaust the time buffer indicated by TSC. For the case when TSC is possibly negative, investigations are necessary to identify measures for increasing TSC.

4.2.2 TSC illustrated

The definitions are better explained with an example. Consider a collision avoidance scenario, shown in Figure 4-1. When two neighboring aircraft become conflicted, i.e. enter a collision course, without control the two aircraft will eventually collide. If the conflict is detected in time, with the control of air traffic controller, for instance, actions will be taken and resolving maneuvers will be applied. At a certain time, the two aircraft will be deconflicted and no longer on a collision course.

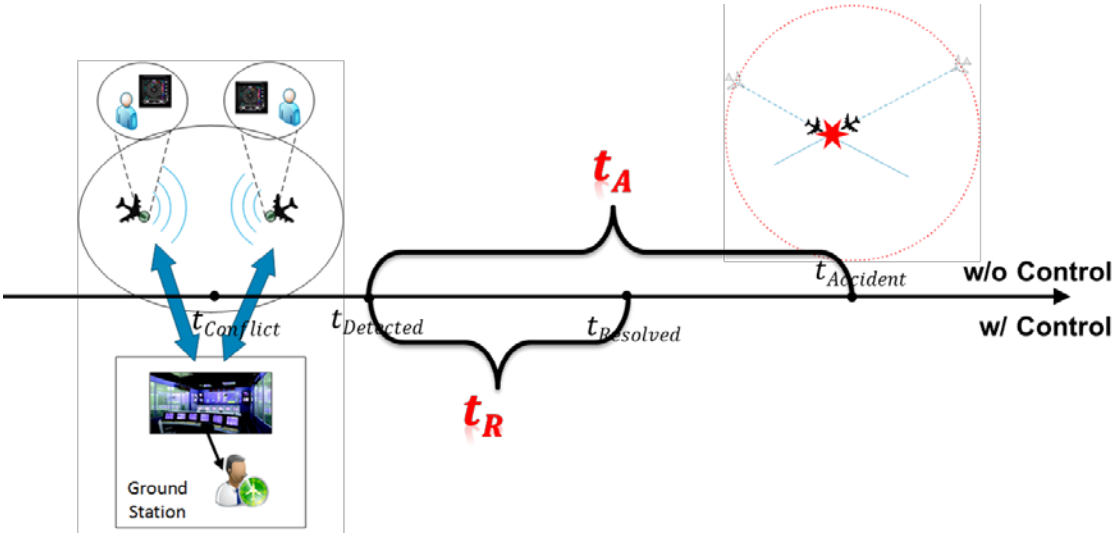


Figure 4-1 Time required and time available

t_A: As is shown in Figure 4-1, at $t_{Conflict}$ on the time axis, two aircraft enter a collision course. Without the control or intervention of ATC, the two aircraft will collide at $t_{Accident}$. Time available for control to take effect as defined is then:

$$t_A = t_{Accident} - t_{Detected} \qquad (4.4)$$

$\boldsymbol{t_R}$**:** The same time instance $t_{Detected}$, when an unsafe situation is detected by ATC, is the initial time instance. In response to the conflict detection, conflict resolution will be made and transmitted to the pilot as control commands for execution. This process will take certain amount of time, and the end of this process, when controller confirms no further action is needed, is marked $t_{Resolved}$. Then

$$t_R = t_{Resolved} - t_{Detected} \qquad (4.5)$$

4.3 Summary

In this chapter, two metrics of control capacity are defined: Probabilistic System Control capacity (PTSC) and Temporal System Control capacity (TSC). PSC is a metric for the system's tolerance of faults or failures in the course of control. And TSC measures system's tolerance of delay before the control objective becomes unattainable. The next chapter further explores the implementation and utilities of the two metrics in quantitative safety assessments.

CHAPTER 5. ASSESS SYSTEM SAFETY WITH CONTROL CAPACITY PART II: METHODS

To use the previously introduced two control capacity metrics, PSC and TSC, in quantitative safety assessments of ATC systems, additional setups are required. This chapter introduces a procedural method assembled to evaluate PSC and TSC, for the purpose of quantitative safety assessments. The method includes three sequential steps: 1) identify safety critical processes, 2) develop control models, and 3) evaluate PSC and TSC.

5.1 Method Overview

Recall the theoretic basis for using control capacity to measure a system's safety performances in CHAPTER 2. Control processes can be characterized and described by states. In a control system's state space, i.e. collection of all possible states, some states are unsafe, such as the accident states. System safety control is to keep the system out of the unsafe state zones. Safety critical processes are state transitions whose failures will result in unsafe event such as an accident. For the system to be safe, all safety critical states need to be controllable. Control capacity is an indicator of system's safety performance when it comes to the control of safety critical processes. In the presence of performance deviations, the higher a system's control capacity over safety critical processes, the safer the system.

For quantitative safety assessments of a given system, introduce a Control capacity Based Safety Assessment Framework (CBSAF), which is an assembly of procedures, methods and metrics required to quantify control capacity, in the control of safety critical processes (see Figure 5-1).

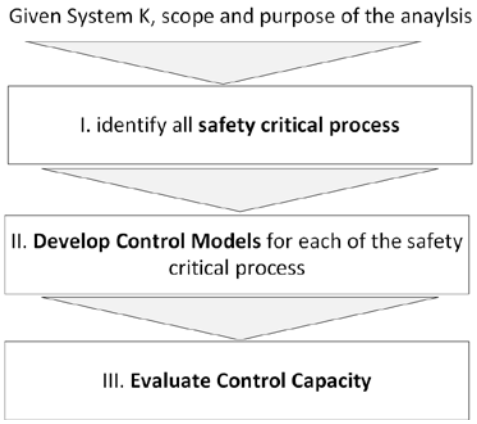


Figure 5-1 CBSAF I: Safety Assessment

Since system control and system safety are correlated through safety critical processes, the safety critical processes must first be identified. Once defined, for each of these processes, the control mechanisms must then be understood and modeled. With the controller and the controlled process known, the control capacity then can be evaluated based on the specificities of the control, i.e. estimation of PSC and TSC.

For each of the step, CBSAF specifies general principles and guidelines needed to carry out the required tasks. The following sections detailed the provided principles and guidelines.

5.2 STAGE I: Identify Safety Critical Processes

For a given control system, there may be multiple safety states, and many safety critical processes. Theoretically, for the system safety to be assured, *all* safety critical processes should be identified and controlled; for a system control capacity to represent the system's safety performance, control over all possible safety critical processes must be examined.

To identify safety critical processes, we use the procedure shown in Figure 5-2.

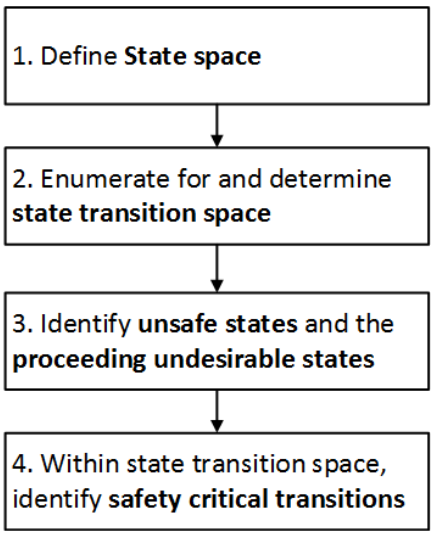


Figure 5-2 Procedure to identify safety critical processes

First, the state space must be defined. Within the state space, all possible state transitions can be then identified through mechanical permutation. Not all states can transit to each other; some state transitions are limited by physical or operational constraints. The exclusion of the "unattainable" transitions requires external determination, e.g. judged by an intelligent agent with understanding of system dynamics. The intelligent agents may be human or computer components, as long as the

criteria are valid. The third step is to identify any proceeding transition that can lead to the unsafe state such as an accident. Precursor states are states that can be detected and escaped from in operation. The controlled escape transition from the precursor state to a safe state is a safety critical process.

The recommended procedure for state space definition is outlined in Figure 5-3. The procedure starts with selecting a set of state variables that characterize the system states. Next, ranges and critical values of the state variables are determined based on the characteristics, time frames and circumstances that are relevant to the safety analysis. The choices of the state variables are dimensions of a “value space”, which is bounded by the extremes of the ranges of these state variables. And the critical values divide the value space into different zones, which is assigned as a state. In control theories, each combination of state variable values defines a system state. For the purpose of identify a finite number of state transitions and thus finite number of safety critical processes, the choices of critical values and the “zones” will limit the number of states to be concerned. For the safety purposes, states that are qualitatively safe, or unsafe are not limited to a particular state variable value, but a small ranges of the state variables, which justifies the use of zones as well.

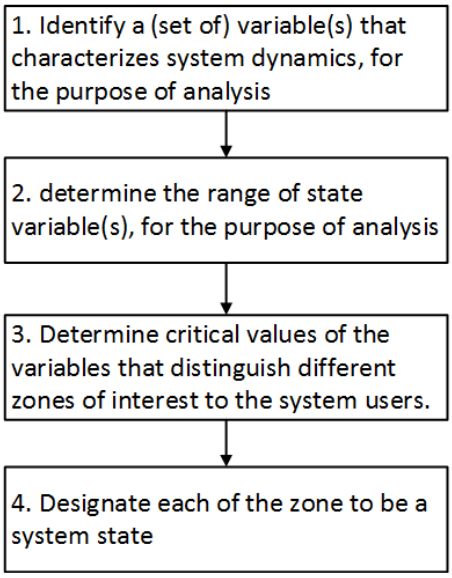


Figure 5-3 Procedure to Define State Space

The overarching procedure of identifying the safety critical processes holds for general safety assessment practice. However, the choices of state variables, ranges and critical values of the state variables, and criteria applied to the classification of safe, unsafety and transitory states may vary for cases with different purposes of analysis and analysts with different levels or sets of empirical experiences and expertise. In other words, there is a level of subjectivity in the execution of the procedures. Yet three general principles are to be followed in the state space definition or used to proof the definitions despite the specific variations.

1) Representativeness: the state should be able to characterize the system for a period of time.
2) Mutual exclusiveness: any two states are mutually exclusive, or that at any given time, the system cannot be in two states simultaneously.

3) Exhaustiveness: the state space should exhaustively represent the system at any time of interests.

5.3 STAGE II: Develop System Control Model

Control mechanisms for the safety critical processes are actualized by system elements, e.g. components and interactions between components. The measuring of PSC and TSC requires specific information about the performance deviations of the specific system elements that involved in the safety control, of a particular safety critical process. The second stage of the method identifies the system elements including components and component interactions, and expresses this control mechanism through the construction of a control model. Control models are the basis for identifying the possible performance deviations, i.e. failure or faults for evaluating PSC and time components for TSC.

As is shown in Figure 5-4, the control model of the national air space ATC system developed by (Haraldsdottir et al., 2001) exhibits the hierarchy and complexity of national ATC. To the right hand side of the control structure, the aircraft is controlled by its pilot with assistance from the automated system, i.e. autopilot. When the number of controlled aircraft increases or the controlled airspace expands, additional control structures are added, for example, sector traffic control, and an outer loop of “facility flow planning”. The outer control loops are strategic, concerning larger airspace capacity, schedule and weather change; the inner control loops are more time sensitive and more critical to the system’s immediate safety.

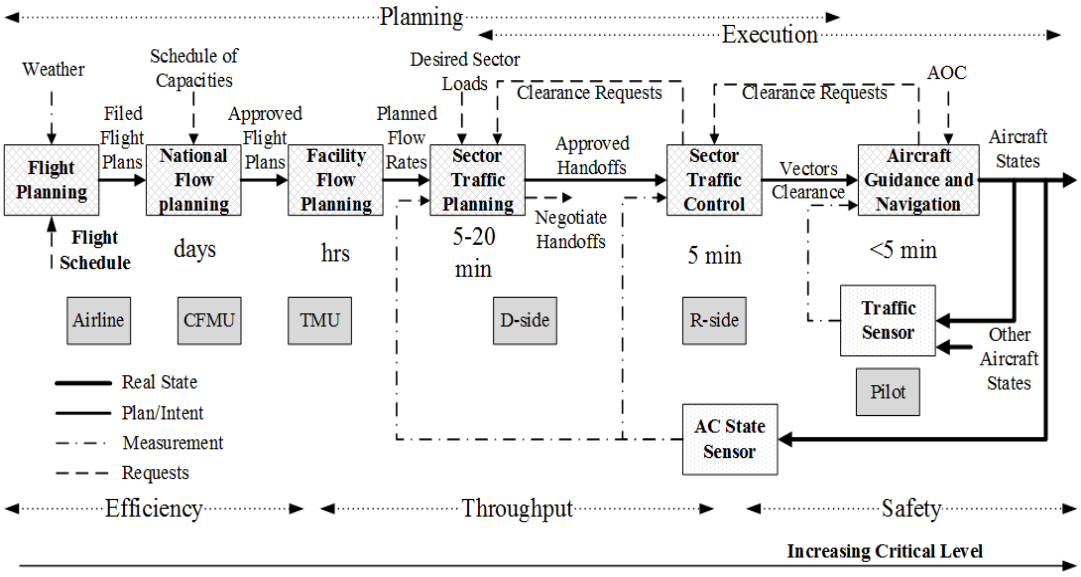


Figure 5-4 ATC operation structure (adapted from (Haraldsdottir et al., 2001))

At each layer of control, the same basic feedback control mechanism is observed. Figure 5-5 shows the basic feedback control loop.

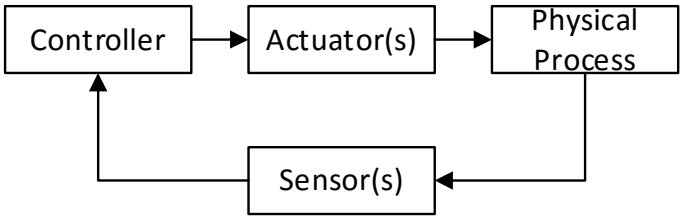


Figure 5-5 Feedback Control Loop

As is shown in Figure 5-5, in feedback control systems, a physical process is typically controlled by a decision maker or a controller. Control decisions are made based on information collected by the sensors and applied to the process through actuators.

Complex systems, particularly socio-technical systems, are better modeled with the adapted feedback control loop structure recommended by (Leveson, 2004), as shown in Figure 5-6.

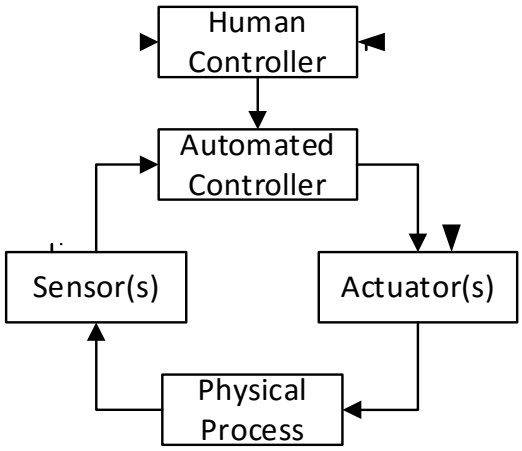


Figure 5-6 Feedback Control Loop Highlighting Human and Automation, adapted from (Leveson, 2004)

As is shown in Figure 5-6 the adapted feedback control loop, both the human and automated controllers have access to the physical process through communication systems, sensors and actuators. However, the human controller will ultimately have higher authority over automation, and typically uses the automated systems to implement the decisions (Leveson, 2004). For instance, between pilot and auto-pilot, the autopilot is controlled by the pilot, and the pilot may use the autopilot to maintain altitude.

The blocks in both models are functions, rather than components, that must be fulfilled by the control system. There is typically no one-to-one mapping from the function to components/system elements. A control function does not require a dedicated component, conversely, a system component may achieve more than one control function.

Redundancy further complicates the generalization of control models. In some cases, more than one component or combination of components can achieve a control

function. Or combinations of some components can achieve multiple functions. The definition of PSC requires a systemic approach.

To address the above two problems, CBSAF proses an adapted control model which takes a top down functional decomposition approach; it emphasizes the completeness of the general control functions rather than specific components. The model also allows redundant components and interactions to be organized in accordance to their functions. This model is shown in Figure 5-7.

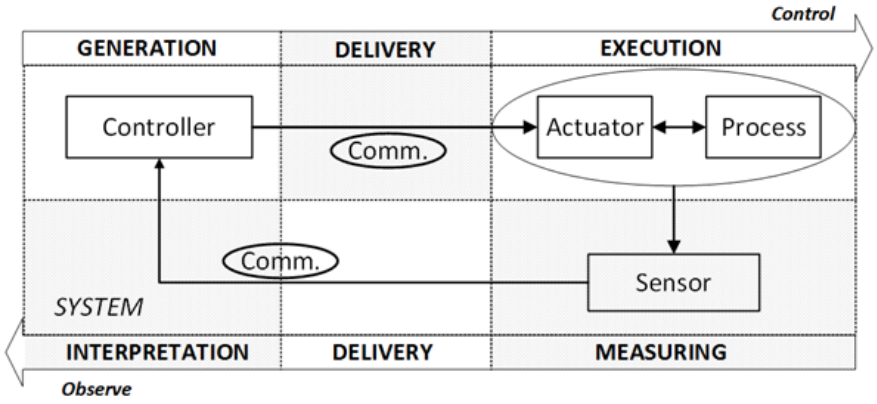


Figure 5-7 Adapted general system control model

As is shown in the adapted general system control model, a control loop is generalized in two sequential functions: observe and control. Observe can be further decomposed into three sub-functions: sensing, delivery and interpretation of the controlled process. Similarly, the control function can also be decomposed into three sub-functions: generating, delivery and execution of control commands. The blocks of sensor, actuator, communication etc. are placed along the observe-control loop as examples of function – system element mappings. The actual system may have different structures from this generic example.

This control model serves as a starting point of constructing a control model for a safety critical system. It provides the first step and guidelines to map out actual system elements in achieving the intermediate functions required for a control loop.

Note that in mapping the system elements to the control functions, as mentioned previously, there is not a one-to-one mapping between each control functions and system components. In some cases, more than one component or combination of components can achieve a control function. Similarly, combinations of some components can achieve multiple functions. In other words, some system elements may need to be placed in between the blocks assigned to the control functions, or several system elements are needed to be in one block. The control model provided in Figure 5-7 serves as a starting point and a generic guideline, and not a requirement.

For the detailed definitions and implications of function, state variable (attributes), and (control) structure, refer to Section 3.2.1.

5.4 STAGE III: Evaluate Control Capacity

As shown in Section 5.1, CBSAF uses both PSC and TSC to measure control capacity considering variations of systems and processes in the control loop. This section details the recommended procedure for estimating PSC and TSC given the control loop specifications from the first two steps.

5.4.1 Evaluation of PSC

PSC can be viewed as a special measure of risk: the risk of losing control. It can be assessed using any of the many readily available tools and techniques for risk assessment. In the case study of this research, Event Tree Analysis (ETA) is adopted to

quantify PSC for its simplicity and intuitiveness. Other probabilistic risk assessment techniques, such as fault tree analysis can also be adapted to estimate PSC of the control loop(s) derived from the first two steps.

5.4.1.1 Overview of Event Tree Analysis

Event tree analysis is a binary form of a decision tree for evaluating various multiple decision paths in a given problem. ETA was first used in WASH-1400, nuclear power plant safety study (Commission, 1975). It was a replacement of FTA to reduce the size of the tree when applied to complex system (subsystems, components, assemblies, software, procedures, environment, and human errors) (Ericson, 2005).

The general process of ETA is given in Table 5-1.

Table 5-1 ETA Process adapted from (Ericson, 2005)

Step	Task	Description
1	Define the system	Examine the system and define the system boundaries, subsystems and interfaces
2	Identify the accident scenarios	Perform a system assessment or hazard analysis to identify the system hazards and accident scenarios existing in the system design
3	Identify the Initiating Events (IE)	Refine the hazard analysis to identify the significant IEs in the accident scenarios; include events such as fire, collision, explosion, pipe break, toxic release, etc.
4	Identify the Pivotal Events (PE)	Identity the safety barriers or countermeasures involved in the particular scenarios that are intended to preclude a mishap
5	Build the event tree diagram	Construct the logical Event Tree Diagram(ETD), starting with IE then the PEs, and completing with the outcome of each path
6	Obtain the failure event probabilities	Obtain or compute the failure probabilities for the PEs on the ETD.
7	Identify the outcome risk	Compute the outcome risk for each path in the ETD
8	Evaluate the outcome risk	Evaluate the outcome risk of each path and determine if the risk is acceptable

9	Recommend corrective actions	If the outcome risk of a path is not acceptable , develop design strategies to change the risk
10	Document ETA	Document the entire ETA process on the ETDs. Update for new information as necessary

A generic event tree diagram is shown in Figure 5-8.

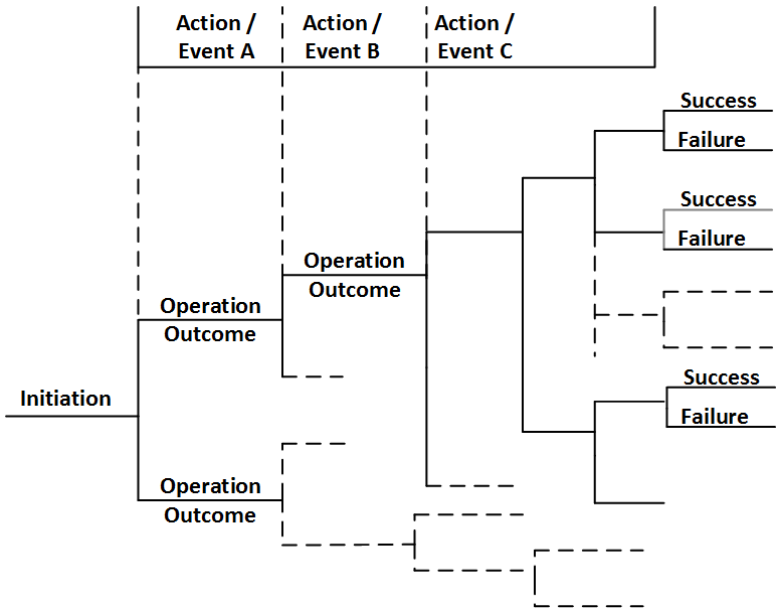


Figure 5-8 Generic Event Tree Diagram

Event trees are one of the most used tools in Probabilistic Risk Assessment (PRA). The advantages to use ETA are highlighted in (Saleh et al., 2010). ETA combines hardware, software, environment and human interactions, can be applied to different levels of details, and models complex system in an understandable manner.

The primary counter argument is that since ETA requires specialty of experts, for analysts of different levels of expertise, it may be prone to errors, and subject to arbitrariness.

5.4.1.2 Recommended procedure to use ETA for PSC estimation

A generic ETA includes three major steps: 1) identification of a hazard list, 2) acquisition of probabilities of each hazard, and 3) identification of chains of hazard event that lead to success/failure.

The first two steps gives the control loops of the control mechanisms over identified safety critical processes. PSC is defined as the probability to retain at least one means of control. Using ETA, the three components are then 1) identification of a hazard list that hampers preservation of control means, 2) acquisition of probabilities associated with each hazard and 3) identification of chains of event that lead to the loss of all control means.

To obtain these three components needed for ETA, CBSAF recommends the following procedure:

1) Based on the control model and the controlled safety critical processes, conduct hazard identifications to as exhaustively as possible, to acquire a hazard list
    a) For each control subfunction, identify its corresponding system elements (both components and component interactions)

    This process should be obvious since the construction of control model, using the generic control model would have mapped the relevant system elements, by placing them in the shaded areas corresponding to each function.

    b) For each of the functions, and their corresponding system elements, identify hazards whose occurrence would result in either loss of function or malfunction.

c) Once completed for all six control sub-functions, the hazard identification is considered complete.

2) From hazard event list in Step 1), enumerate for all chains of events that lead to loss of all control means

a) In each control loop, consider the order of executions of the six control sub-functions, sort the hazards in chronological order along the control loop, and place them on top of the Event Tree Diagram.

b) Assume that each hazard either occurs or it does not occur. Use the event tree diagram to enumerate for all possible combinations of occurring events (or the opposite) that lead to the system's loss of control, or preservation of control.

c) Identify all possible chains of event that lead to success.

3) Estimation of PSC

a) Estimate the probability of each individual hazard identified in 1-c).

b) For each event chain identified in 2-c), calculate the probability of the event chain, using estimation in 3-a).

c) Sum up probabilities estimates of all identified event chain, and this is the PSC estimation.

5.4.1.3 Guidelines for executing the recommended evaluation procedure

Hazard identification: Hazard identification is qualitative and arbitrary. To increase rigor and comprehensiveness to hazard identification, analysts are referred to STPA by

Leveson, where a taxonomy of possible control hazards is summarized (Leveson, 2004). Figure 5-9 shows the categories of control hazards defined in STAMP.

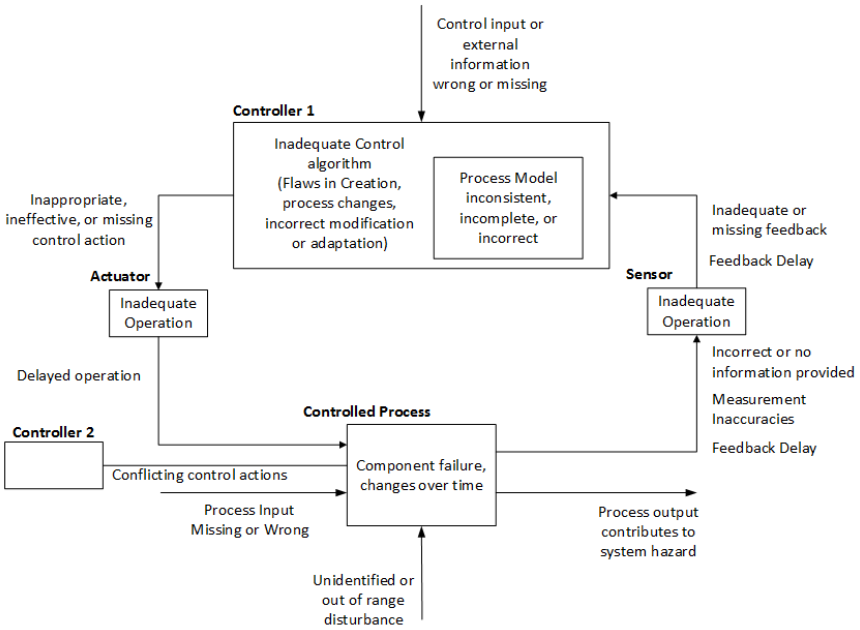


Figure 5-9 Common Hazards in a feedback control loop adapted from (Leveson, 2004)

STPA considers all aspects of control hazards, for the estimation of PSC, not all listed PSC is relevant e.g. unknown hazards, and environment change and delays. Therefore, STPA is recommended as a reference to allow systemic hazard identification activities, but only relevant hazard types are to be considered and included in the event tree analysis.

Probability of individual event: Probability of individual event can be difficult to obtain. In the case studies, values of probability needed are based on judgments and examined by sensitivity tests. Alternatively, these values can be elicited from experts or using performance data of the individual system elements or subsystems.

The utility of PSC lies in comparing systems of different configurations or controls of different processes, and need not to be absolutely correct. In fact, the recommendation

of CBSAF is that, the judgment of probabilities for individual event be considered as a scoring, which preserves the analyst's judgment of the relative likelihood of occurrence for event in the hazard list. A sensitivity test on the choice of values should then be conducted to investigate the validity of the conclusions drawn from the comparison, i.e. which system(s) has higher control capacity.

5.4.2 Evaluation of TSC

TSC is a measure for tolerance of delay, a type of performance deviation. Without performance deviation or delay, the control of a safety critical process should have at least some tolerance of delay, namely time available should always be greater than time required for the system to be safe. During operations, external or internal disruption, disturbance, and degradation will cause the system performance to deviate. TSC is the expected time buffer for system to handle the external performance deviation factors and therefore the higher TSC, and higher this tolerance. TSC measure becomes a safety performance indicator when it comes to the control of safety critical processes.

To measure TSC, the time components are measured under the below three premises.

1) No performance deviation and all temporal values used are nominal/expected;

Since TSC is a measure for tolerance, the expected time buffer to tolerate system disruption, disturbance, and degradations, it is measured with time components at the expected and nominal operation conditions, where there is no performance deviation.

2) Only one control action is needed and taken, and thus one control cycle.

   In other words, the controller only has one shot. Under Assumption 1, this is considered the norm and expected.

This research has a focus on the last minute safety scenarios when it is clear and certain to the controller that without control, the process will degrade into an accident. In this case, the time available to handle the process is often not generously larger than time required, to accommodate interactive controls. Especially for the ATC system which is geographically distributed, for the purpose of safety control, the design to accommodate interactive control means setting detection criteria to allow earlier detection. In the current spatial separation based detection criteria, i.e. at the violation of separation standard, this means large separation standard, which would conflict with the interest to increase airspace capacity.

3) The control cycle starts with an initiating event, and ends with an exit event: the initiating event is a controller's recognition of the event, and the exit event is the controller's confirmation of the attainment of the control objective.

Although it is rare that a system's performance does not deviate, from a designer standpoint, the design has to start from somewhat nominal operational condition. Then taking into accounts the frequent disturbance and disruptions, the designers can determine the appropriate time buffer to handle these frequent delays, either through technologies for early detection (which increases time available), or to reduce time required (e.g. dedicated datalink instead of telecommunications for delivering commands). TSC estimates here are the time buffer that can be estimated by the designer/developers of the system.

The estimation of TSC requires attainment of time available, t_A and time required t_R.

Estimate t_A: For each safety critical process. t_A depends on the specificity of the process. The controlled process is subject to dynamic patterns which can be characterized by a set of time dependent attributes. For example, in the ATC system, the controlled process, air traffic, can be described by the aircraft positions, speed, heading, etc. The attributes of the process are ultimately functions of time. States are certain value(s) of the characterizing attributes such as positions.

The system state evolves over time following the relevant physical, organizational and regulatory principles. The time for the process to evolve from the current state as a new state therefore can be calculated applying system dynamic principles. These principles however vary from case to case. There is no general rule to calculate the t_A for all systems.

Estimate t_R: t_R is determined by steps needed for achieving control and times needed for each step. In estimation of TSC, a generic feedback control mechanism is used as a starting point for constructing a system's control models. From the functional perspective, for a control loop to take effect, six control functions are to actualized consecutively: generating, delivering and execution of control commands, sensing delivery and interpretation of sensed information. As shown in Figure 5-10, the six functions can be unified in that all of them process information. More importantly, information follows through the control loop in one direction.

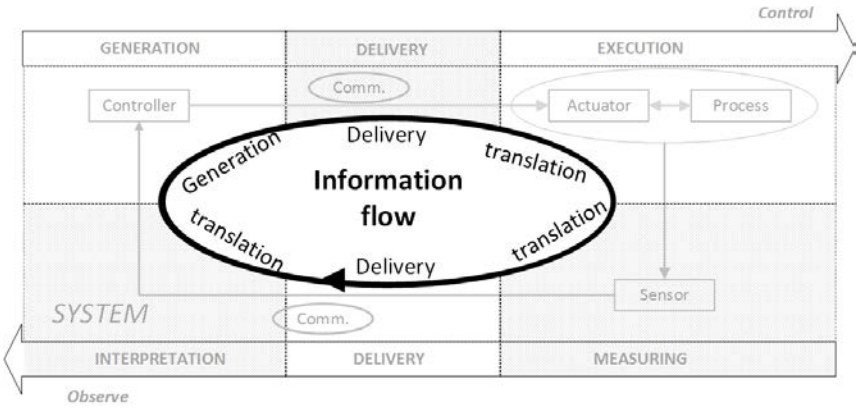


Figure 5-10 Calculating Time Required

In the event of safety control, the control loop will be activated by a controller, upon the controller's identification and recognition of safety threats. As per the feedback control mechanism, from this time instance, a control resolution will be generated, and delivered to the actuators for execution. For the control loop to reactive, then the process after execution will be sensed one more time, and the sensed information will be delivered back to the controller. If the controller interprets the control action to be effective and successful, the control mechanism will be deactivated.

Applying the definition of time required, t_R is estimated by the following equation.

$$t_R = t_{generation} + t_{delivery,control} + t_{execution} + t_{sensing} + t_{delivery,sensing} + t_{interpretation} \quad (5.1)$$

Control functions are actualized by system elements. The construction of control models is intended for mapping the control functions to its corresponding system elements. Actual control systems may have one component to complete more than one function, which requires evaluation of times for the combined steps. Similarly, one control

function may be actualized by multiple system elements, in which case, the time component for this function will require information on times required for each of the system element in the control mission.

The problem is further convoluted by the multiplicity and coupling of control means. In situation when multiple control means work simultaneously, the conservative safety principle asks that the route with the longest time to be used in the TSC evaluation.

The above discussion and the steps for evaluating TSC are summarized by the procedural chart in Figure 5-11, as part of the CBSAF framework.

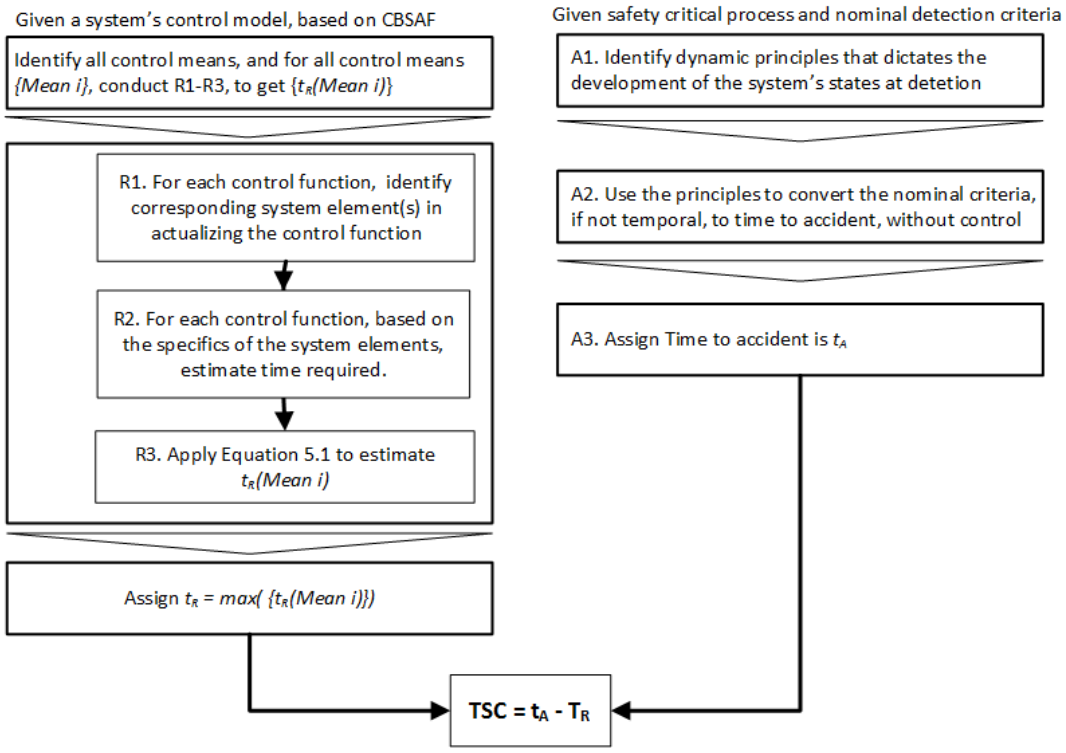


Figure 5-11 CBSAF procedure for TSC evaluation

5.5 Summary

This chapter introduces the theoretical framework, CBSAF, for quantitative safety assessment using control capacity. This framework is consisted of three main steps: I. identify safety critical processes, II. Develop control models and III. Evaluate control capacity. Detailed generalized procedure and guidelines for carrying out each of the procedure is given as well. The following chapters explore the utilities and viability of this framework in the quantitative safety assessments in the ATC systems.

CHAPTER 6. CASE STUDY I: COLLISION AVOIDANCE

A basic air traffic control task is to maintain minimum separation between any two aircraft in the controlled airspace, i.e. 5 nm horizontally, and 1000 feet vertically. A loss of separation conflict is undesirable, since it may be the beginning of a collision. With standard procedures, the ground Air Traffic Controller (ATCo) control is used for strategic conflict resolution to avoid the scenario of collision, while the on board automated collision avoidance System, e.g. Traffic Collision Alert System (TCAS) is used to resolve last minute collision threats.

When a pending collision is detected by both ATCo and TCAS, both the ATCo and TCAS can give instructions to the pilots in making escaping maneuvers. In the case of two instructions were received, standard procedures require pilots to respond only to TCAS, i.e., even if ATCo is instructing them, they must ignore ATCo and respond to TCAS. In the Uberlingen mid-air collision accident, one crew's failure to follow this rule resulted in the fatal accident.

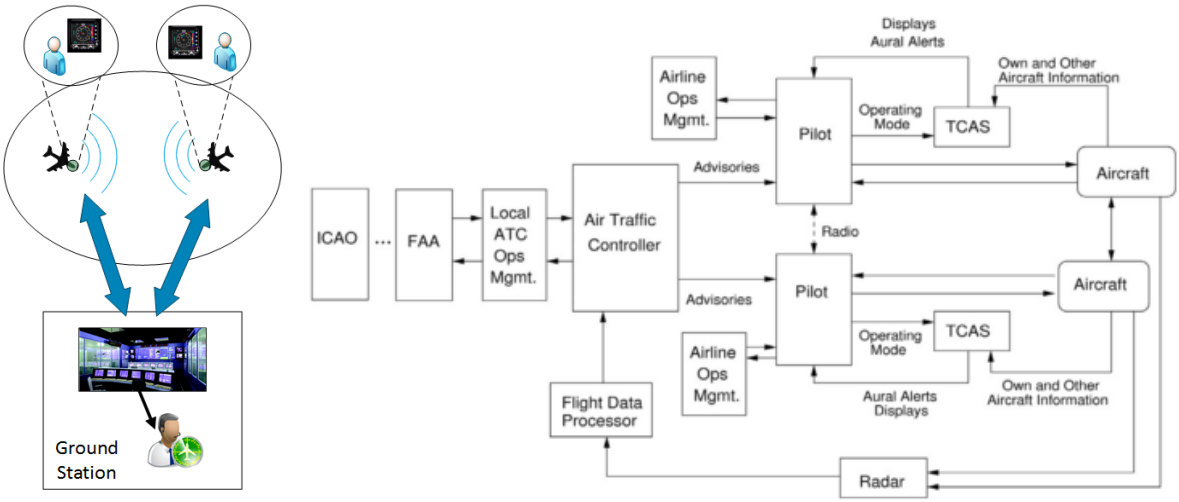
This case study applies CBSAF to the controls of ATCo and TCAS in the control of last minute collision avoidance. To make the comparison, three hypothetical configurations, one with TCAS only, one with ATCo only, and one with both are examined.

The objectives of selecting and setting up the case study is three fold 1) to demonstrate

implementations of the theoretic CBSAF, 2) to uncover implementation issues of the framework, and 3) to preliminarily explore utilities by examining its capability in capturing the specific and expected system behavior observed in the Uberlingen mid-air collision accident.

6.1 Problem description/formulation

The scenario in question regarding the control in collision avoidance is illustrated in Figure 6-1a) and the internal interactions between control system components are illustrated in Figure 6-1b), which is a system diagram using the STAMP model (Leveson, 2015).



a) system illustration b) functional system model adapted form (Leveson, 2015)

Figure 6-1 Control systems of collision avoidance

As is illustrated in Figure 6-1a), in the inner loop, pilots of the two aircraft keep the aircraft on designated routes. The ground controllers monitor the positions of the two aircraft through radar systems. The information is then interpreted by the flight data processor. Should any violation to standard separation occur, the controller will

communicate with the pilots via radio communication and provide solutions to resolve the conflict. TCAS, the onboard automated collision avoidance system, is used as a last minutes collision avoidance; its transponders interrogates with any aircraft (also with TCAS installed) in its close proximity and check for possible collision events. Based on its own predictions, TCAS will determine whether and when to issue advisory to pilots via audio in the cockpit. A Traffic Advisory (TA) is issued at about 35 to 48 seconds to impact, and a Resolution Advisory (RA) about 15 to 35 seconds (DOT, 2011). Normally, the two control means occur at different times following regulations and procedures.

Note the STAMP based model also provides information about the hierarchy external to the control of two aircraft, which is not considered in this research. The case study is scoped to examine the interactions of control means during operations which concerns only lower levels of the system hierarchy.

In the Uberlingen mid-air collision accident, the air traffic controller's involvement and the Russian crew's failure to follow the rules and respond to the TCAS command resulted in an accident. One circumstance in the control is the conflicted instructions from TCAS and the ATCo. For comparison purposes, consider three hypothetical configurations:

- Configuration I: only ATCo is used to command pilots in collision avoidance.
- Configuration II: only TCAS is used for collision avoidance control.
- Configuration III: both TCAS and ATCo are engaged in instructing pilots in collision avoidance.

The three configurations are used to control the same collision avoidance process. Configuration III has both means of control and therefore the interactions should differentiate its control capacity from when only one means of control is used in the control process, i.e. in Configurations I and II. The rest of the chapter details of applying CBSAF to examine the safety performances with the three different system configurations.

6.2 Stage I: Identify safety critical processes

At any time, two aircraft under control are either on the collision course, or not on the collision course. Since the state that aircraft are not on collision course is safe and does not require ATCo to intervene, this state is generalized as X_5.

In the case where two aircraft are on a collision course, without control, the aircraft will travel closer and closer to each other, first violate the separation standards, if not detected by ATCo, then the conflict will reach the threshold and activates TCAS to issue first a TA and then an RA, and if no control is taken, end in a midair collision.

Define state space: A characteristic variable in this process is the time to collision. In fact, this variable is used in the design and development of TCAS II, denoted as τ.

As per (DOT, 2011), τ is defined as:

$$\tau = \frac{Slant\ Range}{Closing\ Speed} \qquad (6.1)$$

Three critical τ values, t_{min}, $\tau(RA)$ and $\tau(TA)$ are used to define four states:

X_1: $\tau \leq t_{min}$; No escape maneuver is available to resolve the conflict, not collided yet.

X_2: $t_{min} < \tau \leq \tau(RA)$; escape maneuver available, and TCAS issues RA.

X_3: $\tau(RA) < \tau \leq \tau(TA)$; escape maneuver available; TCAS issues TA, but not yet RA.

X_4: $\tau > \tau(TA)$; no TA (or RA) is issued by TCAS.

In addition to the no conflict state X_5, the state space concerning the control of collision avoidance between two aircraft is:

$$\{X_1, X_2, X_3, X_4, X_5\}$$

Quantitatively, $\tau(RA)$ is about $15 \sim 35$ seconds and $\tau(TA)$ is about $35 \sim 48$ seconds, depending on the aircraft types, speeds, and headings. In the state space, X_1 is the unsafe/accident state, and should always be avoided. States X_2, X_3 and X_4 are undesirable, since aircraft in these states are on collision course, but not unsafe compared to X_1.

As previously discussed, although choices of states and thus the definition of state space are arbitrary, they must conform to the three general principles: 1) Representativeness, 2) Exclusivity and 3) Exhaustiveness. Additionally, for the purpose of safety assessment, the states must be able to represent the safety critical processes in the system's operations. The defined state space for this case is based on and examined by these requirements.

Enumerate for all possible state transitions: The next step of CBSAF is to enumerate all possible state transitions among the states. The examination may start with a simple permutation of a total of 20 transitions, as shown in Table 6-1.

Table 6-1 State Transition Permutation Table

Initial State	$\rightarrow X_1$	$\rightarrow X_2$	$\rightarrow X_3$	$\rightarrow X_4$	$\rightarrow X_5$
X_1	--				
X_2		--			**
X_3			--		
X_4				--	
X_5					--

The physical constraints will eliminate most transitions e.g. a collision state cannot transit to any other states. Manual examination, based on the physical principles yields 8 possible state transitions: $X_2 \rightarrow X_1, X_3 \rightarrow X_2,\ X_4 \rightarrow X_3, X_5 \rightarrow X_4,\ X_2 \rightarrow X_3\ , X_3 \rightarrow X_4,$ $X_2 \rightarrow X_5, X_3 \rightarrow X_5$ and $X_4 \rightarrow X_5$.

Therefore the state transition space is: $\{X_2 \rightarrow X_1, X_3 \rightarrow X_2, X_4 \rightarrow X_3, X_5 \rightarrow X_4,\ X_2 \rightarrow X_3\ , X_3 \rightarrow X_4, X_2 \rightarrow X_5, X_3 \rightarrow X_5\ X_4 \rightarrow X_5\}$.

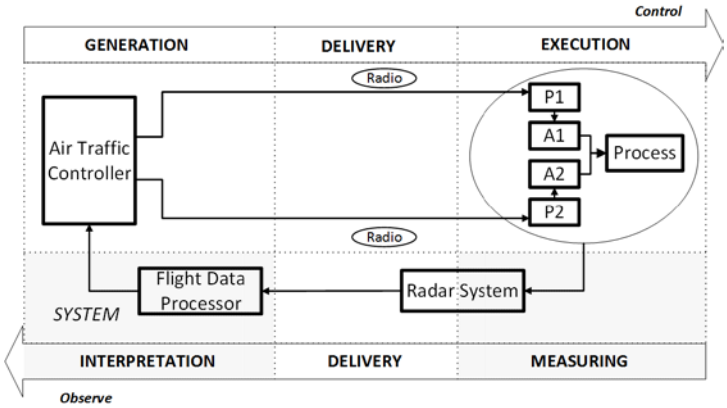
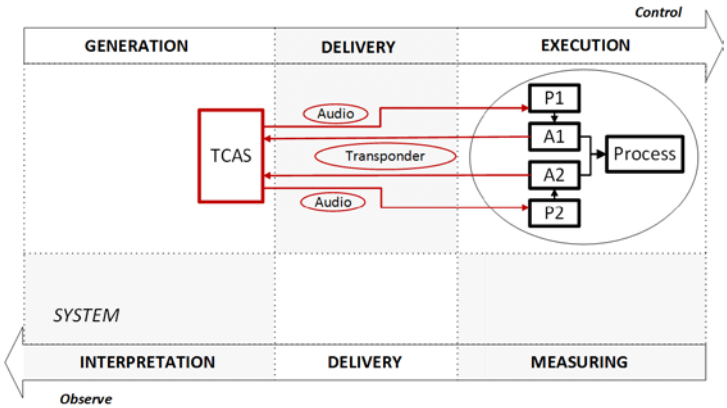
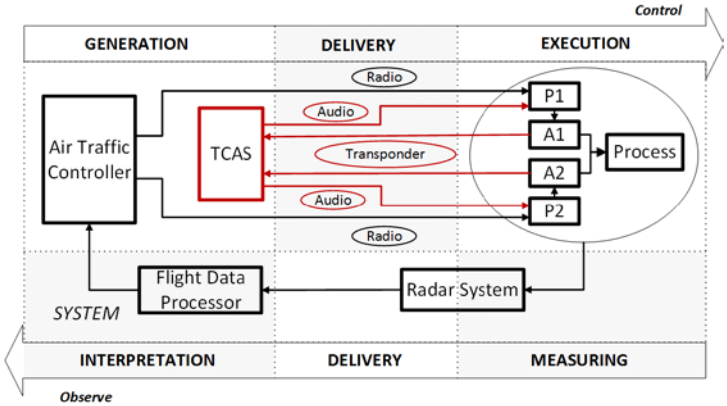
Determine the safety critical processes: Safety critical processes by definition constitute a subset of the transition space. In general, any state other than X_5 is undesirable. Since the scope of problem in question is about collision avoidance, which necessitates the states to at least represent aircraft on collision course, only X_1 is considered to be the unsafe state in the case study.

Based on the definition of safety critical process, that without control, will degrade into an accident state. Enumerate through all possible state transitions and apply this definition. The safety critical process is determined to be p_{cr}: $X_2 \rightarrow X_5$, which is marked "**" in Table 6-1.

6.3 Stage II: System control models

Control Models: The safety critical process, p_{cr}: $X_2 \rightarrow X_5$, is the control of two aircraft from shortly after the RA is issued to when the conflict is resolved. Consider control with the three different system configurations. Based on the six basic functions of a feedback control loop discussed previously, system control model for the three configurations are developed. The individual components and specific communication interactions realizing each of the six basic control functions in the baseline model (See Figure 5-7) are identified sequentially. The final control models of the three configurations are illustrated in Table 6-2.

Table 6-2 Three Control Configurations of Collision Avoidance Air Traffic Control

Configuration	System functional Model
Configuration I: With ATCo only	
Configuration II: Without TCAS only	
Configuration III: With both ATCo and TCAS	

6.4 Stage III: Quantify System Control capacity

6.4.1 Probabilistic System Control capacity

Recall PSC definition, for a given system *K* and process p

$$PSC(K,p) = P(\cup F_i) \text{ or } PSC(K,p) = 1 - P(\cap \overline{F_i}).$$

Where, $\cup F_i$ is symbolic denotation of the event that at least one means of control is preserved, and $\cap \overline{F_i}$ is symbolic denotation of the event that all means of control are lost. First part of the method require control failure/faults to be identified and listed. The control models given in Table 6-2 are examined for this purpose. As is shown in Table 6-3, the identification of fault and failure events start from the six control function to the most left. It then refers to the system elements that actualize each of the functions, in each control means. The portion of control elements then are used to identify the specific faults or failure events. A list of hazardous control hazard events is determined as an outcome of this step, which is listed in Column 6 in Table 6-3.

Also listed are probabilities determined, which are representation of assumed performance deviations. For comparison purposes, these probabilities are subjective scores chosen by the analyst, reflecting the analyst's expectation and judgement on the likelihood and importance of each performance deviation. The values of the probabilities should also be considered assumption on performance deviation. Control capacity is defined as the extent to withstand performance deviation. Namely, between two control systems and given these performance deviations, determine which one will have the higher tolerance.

Table 6-3 Hazardous control events and probabilities

Order	Fun.	Means I	Means II	Type	Event Description	~P
1	Sensing	Radar System	TCAS transponder	H	Radar Fails	0.2
2				S	TCAS fails	0.1
3				H	Radar fails to determine states	0.2
4				S	TCAS fails to determine states	0.1
5	Interp.	FDP	TCAS computer Unit	H	FDP fails	0.2
6				S	FDP fails to reflect states	0.2
7		ATCo		H	ATCo fails	0.1
8				S	ATCo fails to interpret FDF inputs	0.2
9	Gen.			H	TCAS fails to generate resolution	0.1
10				S	ATCo fails to generate commands	0.3
11	Delivery	ATCo → P1/P2 through radio	TCAS → P1/P2 through audio	S	ATCo fails to deliver commands to P1	0.3
12				S	ATCo fails to deliver commands to P2	0.3
13				S	TCAS fails to deliver resolution to P1	0.1
14				S	TCAS fails to deliver resolution to P2	0.1
15				H	P1 fails	0.1
16				H	P2 fails	0.1
17	Execution	ATCo commands	TCAS resolution	S	Compare TCAS resolution with ATCo commands: contradict or otherwise	0.5
18		$P1 \rightarrow A1/$ $P2 \rightarrow A2$	$P1 \rightarrow A1/$ $P2 \rightarrow A2$	H	A1 fails	0.1
19				H	A2 fails	0.1
20.1				S	P1 controls A1 following ATCo, given that TCAS contradicts ATCo	0.5
21.1				S	P2 controls A2 following ATCo, given that TCAS contradicts ATCo	0.5
20.2				S	P1 controls A1 following ATCo, given that TCAS does not contradict ATCo	0.5
21.2				S	P2 controls A2 following ATCo, given that TCAS does not contradict ATCo	0.5
22		A1/A2	A/A2	S	Actuators fail to reach desired states	0.2

H = Hardware, S = Software, FDP = Flight Data Processor, Gen = Generation, ATCo = Air Traffic Controller, A = Aircraft and P = Pilot.

Also listed in Table 6-3 are the two possible means of control in the collision avoidance ATC, types of failure, and assumed probabilities for each event. The results are obtained following CBSAF.

Based on the system's diagrams from Figure 6-1a), Event Tree Analysis (ETA) is then used to enumerate for chains of events that lead to the loss of means of control and quantify the probabilities of the occurrence of these event chains. The resulting event trees for configurations I and II are illustrated in Figure 6-2 and Figure 6-3.

ETA systemically enumerates combinations of the outcomes of the hazard events that lead to obtaining its control objective. Note in Configuration III, controls from ATC and TCAS are analyzed as one process and therefore the interactions between the redundant functions during the control process are accounted for.

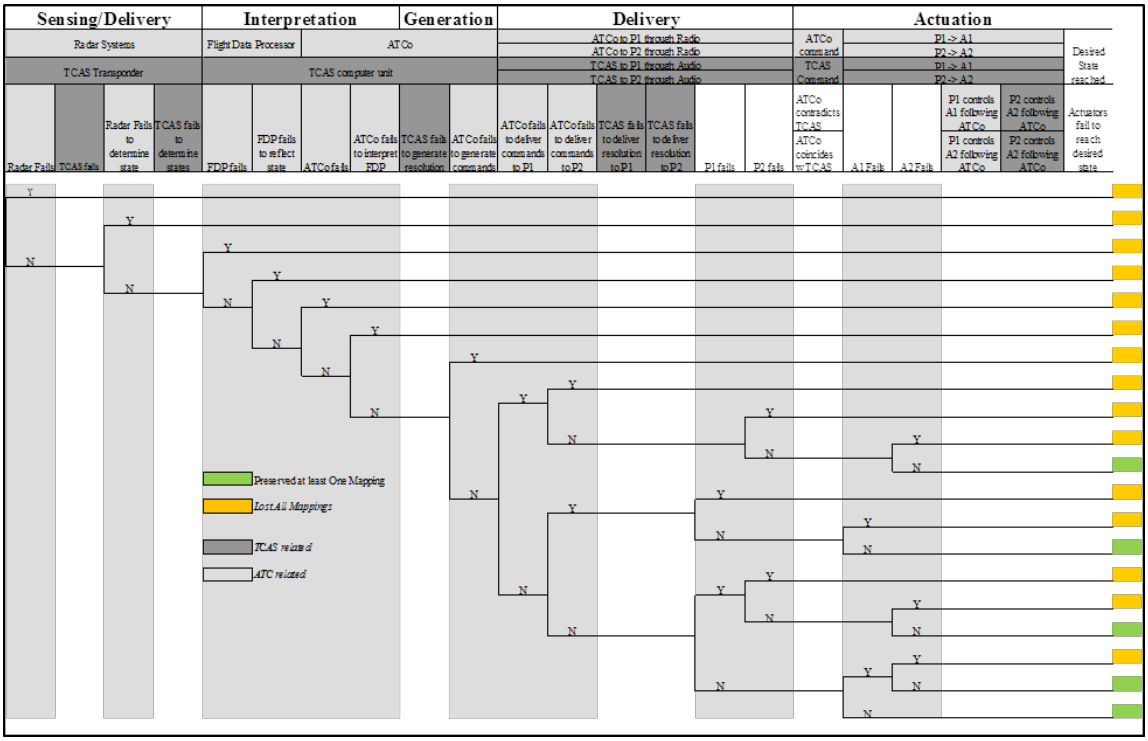


Figure 6-2 Event tree with Configuration I

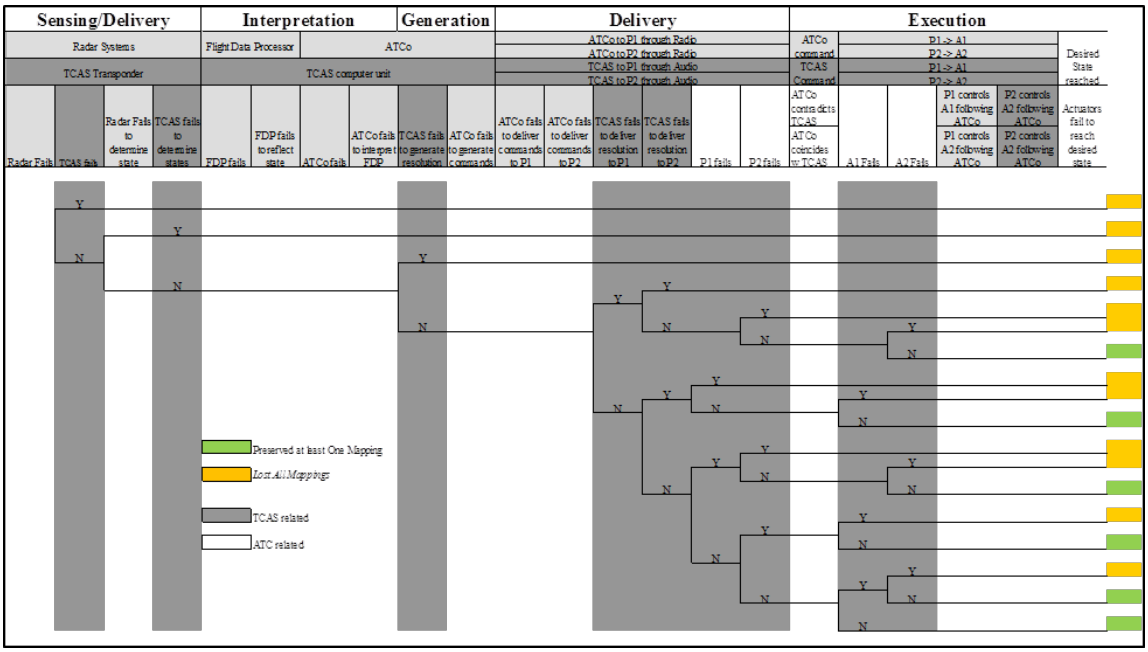
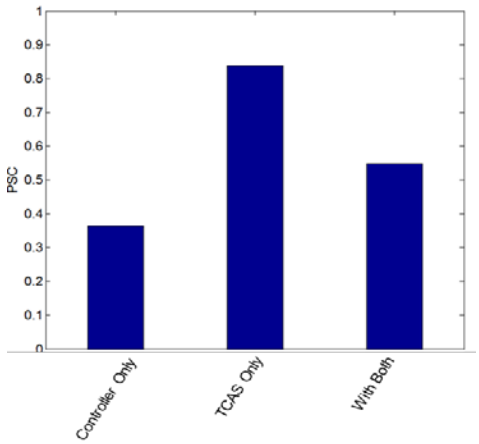


Figure 6-3 Event tree with Configuration II

Applying the probability values for each failure event from Table 6-3 to their ETA gives the estimated PSC of systems with Configurations I, II, and III to be 0.36 0.84 and 0.55 (See Table 6-4).

Table 6-4 Estimated PSC for Configs. I, II and III

System	Config. I	Config. II	Config. III	Change
PSC	0.36	0.84	0.55	- 34.5%



6.4.2 Temporal System Control capacity

6.4.2.1 Time required t_R

For each cycle of feedback control over $p_{cr}: X_2 \to X_3$, the control system will step through the six control sub-functions required for the control loop: generation, delivery, and execution of control commands, sensing, delivery and interpretation of sensed information and TSC is the sum of times used at each step.

6.4.2.1.1 Means of control I: "Radar-ATC-Pilot-Aircraft"

Generation: If a conflict is detected, the air traffic controller will responsively take action to rectify the situation and resolve the conflict. The generation of control commands is determined by experiences of the traffic controller, and complexity of the situation, but the time consumption on this step will be bounded. In collision avoidance, the situation is best resolved by having one aircraft climb up and the other descend. Assuming in this step, the controller only needs to verify if there is any nearby traffic invaliding this solution, based on the radar display of the traffic in the controlled region. Estimate 1-2 seconds to make such verification, and hence $t_{generation} \in (1s, 2s)$

Delivery: Once a decision is made, the control resolution will be communicated to the pilot through radio. The time required is the nominal time to pass a maneuver instruction from the controller to the pilot, using ATC phraseology. This process includes sending of the message and confirmation of the pilot to confirm the pilot's reception of the command by repeating the message back to the controller. Without performance deviation, this process is estimated at 2-5 seconds, thus $t_{delivery} \approx 5s$.

Execution: Upon reception of instructions, the pilots are expected to maneuver the aircraft as instructed. In the last minute collision avoidance, the standard resolution is to have one aircraft climb to a higher altitude and the other descend, e.g. by 1000 feet. A typical commercial jet takes about 3-5 second to resolve the altitude conflict. Therefore, estimate $t_{execution} \in (3s, 5s)$

Sensing and delivery: ATC around the world uses the Air Traffic Control Radar Beacon System (ATCRBS) to locate and identify aircraft. The components and working principles of ATCRBS are illustrated in Figure 6-4.

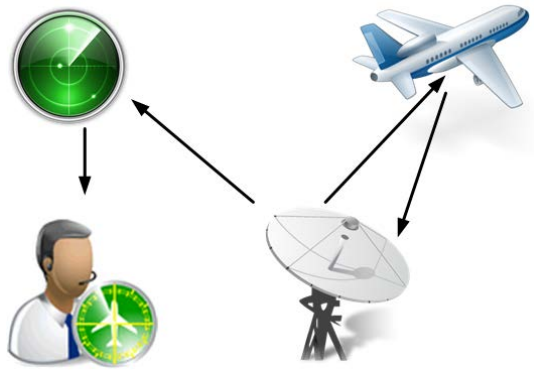


Figure 6-4 Radar System

As shown, the time needed for radar to detect aircraft position comprises two parts: 1) interrogations and 2) sweep rate. The interrogation is the transmitting and receiving of pulses (radio frequency electromagnetic signals). Since electromagnetic signals travel at the speed of light, for the range radar system is used in ATC, it only takes between 40 ms and 100 ms for the signal to travel in most traffic control scenarios. The sweeping rate on the other hand, will contribute to seconds of delay. For example, the Honeywell

RDR 2100 takes 3-4 seconds to perform a 90^{o} scan. Based on this information, the total time estimated by for sensing and delivery is between 0.1 seconds to 5 seconds.

Interpretation: The last step is for the controller to confirm that the control has effectively resolved the safety threat. In this case, information of the two involved aircraft position, speed, and headings will be displayed on the radar screen. The controller will project based these information if the two aircraft are relieved of collision threat. For the resolution of altitude change, this should be straightforward, that if radar shows that the two aircraft are on different altitudes, the safety control loop will be deactivated. Estimate that this interpretation will take 1 to 2 seconds. Hence, $t_{interpretation} \in \{1s, 2s)$.

6.4.2.1.2 Means of control II: "TCAS-pilot-Aircraft"

The application of TSC definition to the control of TCAS is not as straightforward, because the controller in this model is an automated system, rather than a human. The same principles apply however, that until the TCAS has detected a threat, the control loop will not be activated. Similarly, until the two aircraft's physical attributes indicate no further threat, the TCAS to pilot interaction will be active, i.e. audio advisory will continue to be issued.

TCAS mechanism: One complication in evaluating TSC in this means of control however is that the TCAS is one unit that accomplish multiple functions. To examine for the control functional components, the TCAS mechanism is shown in Figure 6-5.

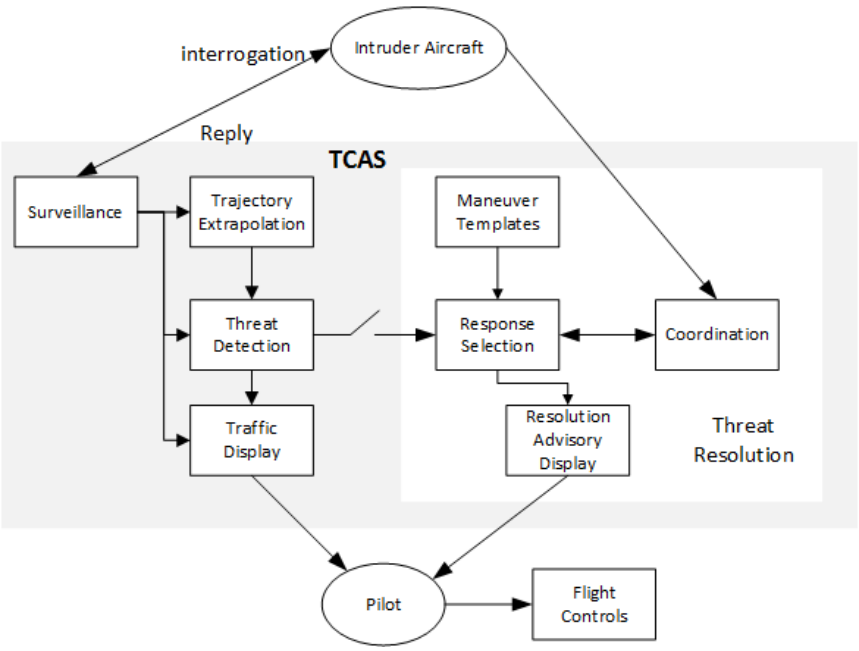


Figure 6-5 TCAS mechanism (Kuchar and Drumm, 2007)

As is shown in Figure 6-5, control functions sensing and delivery are achieved by the surveillance transponder unit, which interrogates any “intruder” aircraft in its proximity. Information of the surveillance unit is then fed to a unit that extrapolates projected trajectories of the “self” aircraft and the “intruder” aircraft for detection of any imminent safety threats. Should any threat be detected, the threat resolution unit will generate resolution advisories i.e. RA or TA, and announce them to the pilot at appropriate times.

Generation: The resolution advisor is pulled from a template database. Estimate that this step takes no more than 0.5 second and as low as 0.1 second. Therefore, $t_{generation} \in (0.1s,\ 0.5s)$.

Delivery: Advisories of TCAS are communicated to the pilots through cockpit audio. The delivery completes when pilots accept the advisory and form intent to follow TCAS instruction. This process is estimated to be $t_{delivery} \approx 2s$.

Execution: Resolution advisories of TCAS to resolve a pending collision has only two types: climb or descend. Similar to the case with the other control means, the execution of this escape maneuver is estimated at 3-5 seconds. Hence, $t_{execution} \in (3s, 5s)$.

Sensing, delivery and interpretation: Since these functions are actualized by the transponder unit, by TCAS, this process is given a time estimation of $0.1 - 0.5\ seconds$. Thus, $(t_{sensing} + t_{delivery} + t_{interpretation}) \in (0.1s, 0.5s)$

6.4.2.1.3 t_R Summarized

The above discussion on the time components for evaluation of TSC is summarized in Table 6-5.

Table 6-5 Times needed for each step

Means of control	ATC to A1 or A2		TCAS to A1 or A2	
Generation	Air Traffic Controller	1–2s	TCAS: Pull from template	0.1–0.5s
Delivery	Radio Communication	5s	Cockpit Audio	2s
Execution	Maneuver of aircraft	3–5s	Maneuver of aircraft	3–5s
Sensing	Radar	0.1–5s	TCAS: Transponder interrogation and prediction computing	0.1–0.5s
Delivery	Radar to ground			
Interpretation	Flight Data Processor to controller	1–2s		
t_R	10.1s–19s		5.2s–8s	

Therefore, for TCAS, the time needed for completing one iteration is

$$t_{L,TCAS} = t_{TCAS,generation} + t_{audio} + t_{pilot\ \&\ aircraft} + t_{TCAS,Observe} \in (5.2\ s, 8\ s)$$

For ATC, the time needed for completing one iteration is

$$t_{L,ATC} = t_{controller} + t_{radio\ communication} + t_{pilot\ \&\ aircraft} + t_{radar} + t_{data\ processor} \in (10.1s, 19s)$$

6.4.2.2 Time available t_A

Assume ATCo will be alerted at 5 nm slant range between the two aircraft; and pilot will be alerted by ATC or TCAS whichever is earlier. Typical commercial aircraft cruising speeds are about 450 kts.

At the worst case, when two aircraft fly head on towards each other, at the time when ATCo is alerted, time available is

$$t_A(ATCo) = \tau = \frac{Slant\ Range}{Closing\ Speed} = \frac{5}{450 + 450} \times 3600s \approx 20\ s$$

The state variable selected and the literature on TCAS already gives time to conflict information about X_2. As defined X_2 is between the RA is issued and when there is no escape. This time as per TCAS II is about 19s to 35s to conflict. The estimated time available also falls into the range of time.

6.4.2.3 TSC

Given the above estimations on t_R and t_A, the TSC for all three system configurations are calculated based on Eq 4.3 and summarized in Table 6-5.

Table 6-6 Temporal System Control capacity for $X_2 \rightarrow X_5$

	With ATCo only	With TCAS only	With Both
t_R	$(10.1s, 19s)$	$(5.2\ s, 8\ s)$	$(5.2s, 19s)$
t_A	$20s$	$20s$	$20s$
TSC	$(1s, 9.9s)$	$(12s, 14.8s)$	$(1s, 14.8s)$

6.5 Result Analysis and Sensitivity Test

6.5.1 Comparison

PSC: As shown in Table 6-4, there is a significant difference between the PSC values of the control of Process $X_2 \rightarrow X_0$ for the systems with Configurations I and II. The probability values of events in Table 6-3 are intentionally inflated to augment the differences. The probability values are treated as assumptions on performance deviation, and scored by analysts as a representation of analyst's judgements on the likelihood and importance of each failure/fault hazard.

The first observation from the result is that the difference of Configurations I and II, 0.48, compared to that between Configurations II and III 0.29, is much larger. There is an indicator that the function of TCAS in increasing control capacity is significant.

Also noted is a decrease of 34.5% PSC is observed from Configuration II to Configuration III. Since Configuration III has both means of control, and Configuration II has only the air traffic controller, this decrease in control capacity suggested the addition of air traffic controller means has affected the control effectiveness negatively and by a significant amount. Since the control of collision avoidance is safety critical, this control capacity is

a measure to system safety performance. Therefore the conclusion drawn from this comparison is that the introduction of ATCo in the collision avoidance process will result in a less safe control system.

The Überlingen mid-air collision accident is the only event where the system with Configuration III was by chance the case. The conclusion from this case study is aligned with the observation in the Überlingen accident that it is less safe to have both air traffic controller and TCAS control a pending collision.

TSC: The results of TSC evaluations show that there is a higher tolerance to delay for Configuration II with TCAS only than Configuration III with ATCo only. The estimate for Configuration I shows a minimum of 1 s TSC and for configuration II a minimum of 12 s. In ATC systems, delays are frequent and common. For a system with a tolerance of close to 0 s indicates highly likely scenarios where there may not be enough time to respond to a safety situation. In contrast, TSC estimate for Configuration II indicates a bigger temporal protection zone to absorb any delays. Since TCAS was designed and configured with temporal criteria and the ATCo is not meant for collision avoidance, such results are within expectation.

The estimates of TSC of Configuration III are based on the worst and best case scenario assumptions, taking the union of TSC ranges of Configuration I and Configuration II rather than from testing the actual execution of the controls. To acquire estimates that account for the interactions between the different control means, alternative approaches are recommended such as through experimenting on the concurrent execution of the two means of control.

Additionally, the time components used for evaluation TSC for the three configurations are based on rough estimates. This is due to the demonstration nature of the case study and limited resources on experimental and actual data. For more meaningful comparisons between the different configurations, estimations of the time components are recommended to refer to more reliable sources and references.

6.5.2 Sensitivity Test on PSC evaluation

In the evaluation of PSC, probability values of individual hazards are based on judgements, which reflect the analysts evaluation of each of the hazard's likelihood of occurrence and importance should its performance deviates. This section is set up to test impacts of subjectivity in these judgements, by choosing a wide range of probabilities of occurrences for the hazard list and check if the same qualitative comparison results will be reached.

Values used in Table 6-3 are varied in two ways: 1) a percentage uncertainty is attached to each probability, and 2) a single sided uncertainty is given to each probability value. The new ranges of probability values are given the

Table 6-7 Tests Details

Test	ID of Hazards. applied	Original P	Tested P range	Select criteria
1	$\forall i \in [1,22]$	P_i	$P_i(1 \mp \delta_1\%)$	*Uniform distribution*
2	$\forall i \in [1,22]$	P_i	$P_i + (0, \delta_2)$	*Uniform distribution*

For Test 1, four $\delta_1\%$ values are used: 10%, 30%, 50% and 90%. For 5000 runs, the ranges of PSC for Configurations I-III are shown in Figure 6-6 to Figure 6-9.

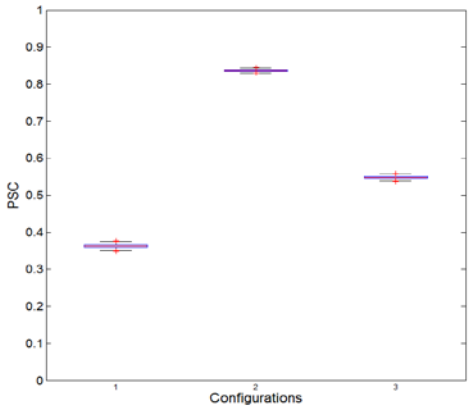


Figure 6-6 PSC distribution for $\delta_1 = 10$

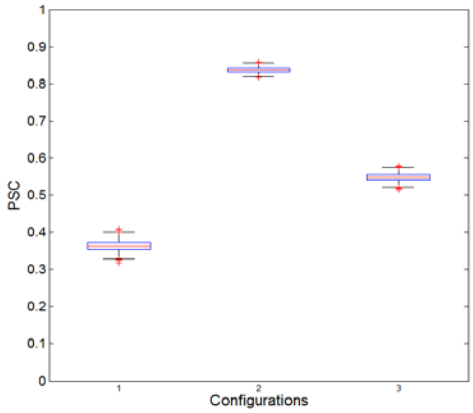


Figure 6-7 PSC distribution for $\delta_1 = 30$

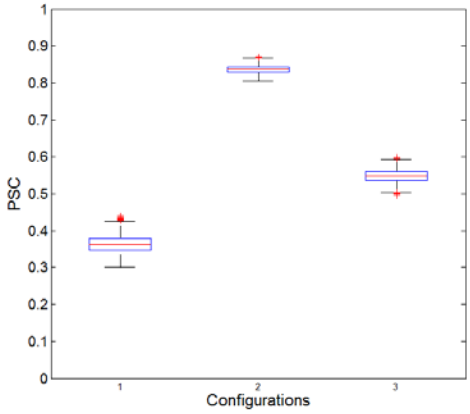


Figure 6-8 PSC distribution for $\delta_1 = 50$

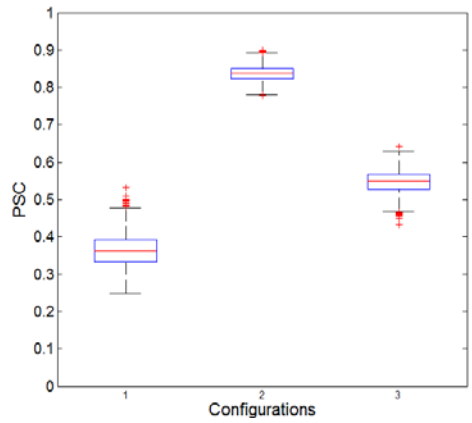


Figure 6-9 PSC distribution for $\delta_1 = 90$

As can be seen in Figure 6-6 to Figure 6-9, the increase of level of uncertainty increases the ranges of PSC each of the configuration can be. However, the uncertainty introduced was shown to be insignificant in that it does not change the comparative orders of PSC for the three different configurations. For the 5000 combinations of slightly different event probabilities at the designated $\delta_1\% = 10\%$ uncertainty levels, the average remains that: $\overline{PSC(I, p_{cr})} \approx 0.38$, $\overline{PSC(II, p_{cr})} \approx 0.84$, and $\overline{PSC(III, p_{cr})} \approx 0$.

Because the values of probability of occurrence selected for each hazard event are different in order to represent the relative likelihoods, the absolute variation applied to each probability of occurrence is different for each event as well.

For Test 2, consider if uncertainty intervals are the same for all $P(E)$. For the cases of $\delta_2 = 0.1, 0.2, 0.3\ and\ 0.4$, distributions of PSC are illustrated in Figure 6-10 to Figure 6-13.

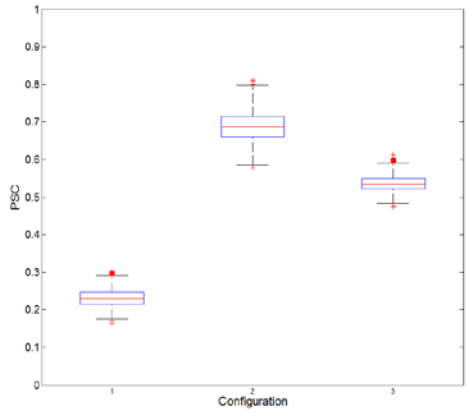


Figure 6-10 PSC distribution for $\delta_2 = 0.1$

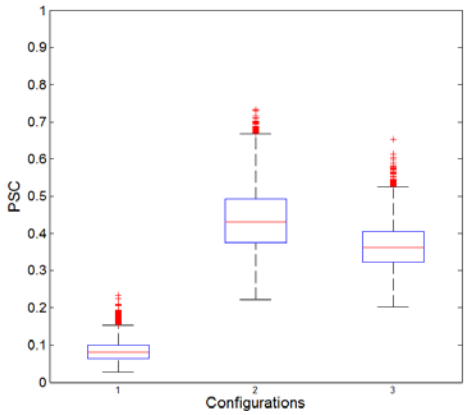


Figure 6-11 PSC distribution for $\delta_2 = 0.3$

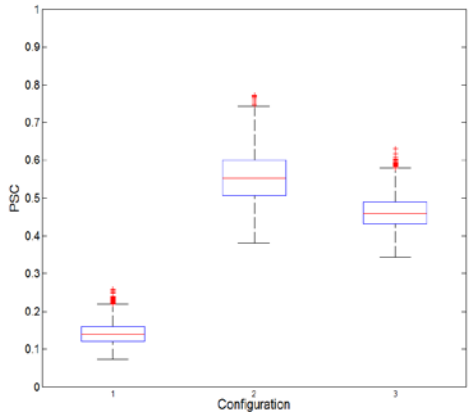


Figure 6-12 PSC distribution for $\delta_2 = 0.2$

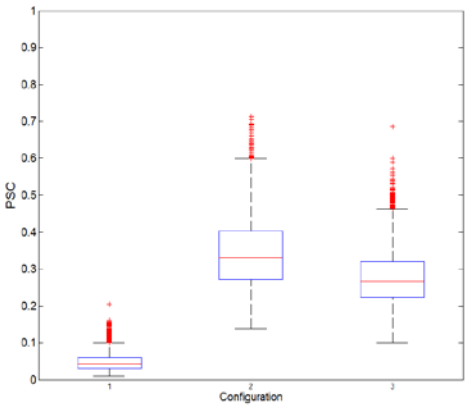


Figure 6-13 PSC distribution for $\delta_2 = 0.4$

As is in Figure 6-10 to Figure 6-13, since the values of occurrence probabilities increase as δ_2 increases, overall the PSC will grow smaller with δ_2. Variation of PSC also increases

as the uncertainty level increases. For the PSC(II) to be smaller than PSC(III), the P values need to be above 20%. This sensitivity test provides evidence that the comparison conclusions from Section 6.4 hold and are robust to the variations of the judgements on the probabilities of occurrence for the elementary hazard events.

6.6 Summary

This chapter applies the CBSAF to the study of en route collision avoidance, particularly to capture the incoordination between the controller and the TCAS observed in the Uberlingen accident. Three system configurations are considered for the safety critical process: $X_2 \rightarrow X_5$, where X_2 is a precursor state, without control, would degrade to the unsafe state X_1, an accident. PSC and TSC are measured based on the control models, and compared among three hypothetical configurations. The results indicate that the involvement of ATCo over $X_2 \rightarrow X_5$ in fact compromises the system's safety performance. This is aligned with observations in the Uberlingen accident.

This case study also uncovered a number of feasibility to ATC safety assessment problems. The discussion on the methods limitation is detailed in the Chapter 8.

CHAPTER 7. CASE STUDY II: RUNWAY INCURSION

7.1 Background and Problem statement

7.1.1 Background

The International Civil Aviation Organization (ICAO) defines runway incursion as “any occurrence at an aerodrome involving the incorrect presence of an aircraft, vehicle or person on the protected area of a surface designated for landing and takeoff of aircraft”. Runway safety has been on the NTSB's annual list of "Most Wanted Improvements" since 1990. Yet the number of incursions reported in the U. S. rose from 186 in 1993 to 431 in 2000, an increase of 132% (Jones, 2002). It is also reported that over 80% of pilot caused runway incursions occur during taxi to the departure runway.

Ground controllers, runway controllers, pilots, stop bars, and automated systems are part of the ATC system to control and avoid runway incursion accidents. In the US, two automated systems: Airport Surface Detection Equipment, Model X (ASDE-X) and the Airport Movement Area Safety System (AMASS) are used to alert air traffic controllers to the potential for a runway incursion. In Europe, similar automation systems termed Runway Incursion Alert System (RIAS) are used. The automated systems are designed to provide an alert 15 seconds before the aircraft reach the conflict point (Stroeve et al., 2013).

Being able to understand and assess the safety of runway operations beforehand is essential to the assurance and improvement of runway safety. A Multi Agent Dynamic Risk Modeling (MA-DRM) based quantitative safety assessment method was developed in (Blom et al., 2006, Stroeve et al., 2009, Stroeve et al., 2013), where the runway incursion safety assessment for systems with and without RIAS is conducted.

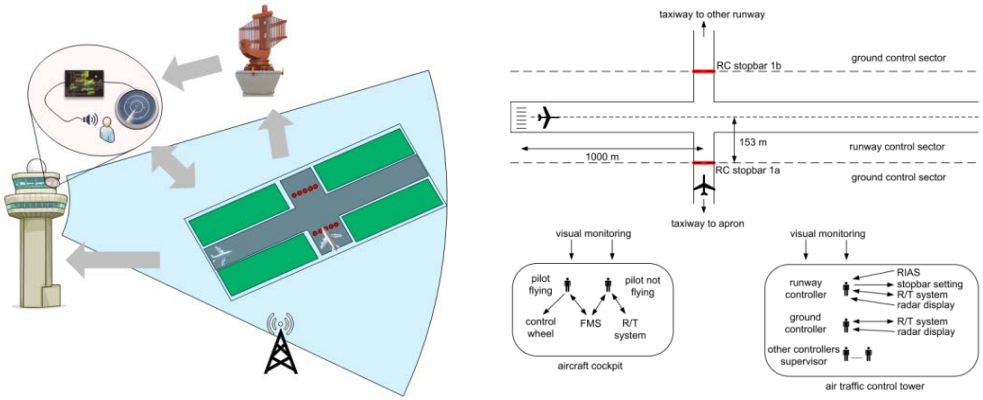
This chapter compares the CBSAF with the MA-DRM method in the case of the runway incursion quantitative safety assessments. Similar setup and terminology as the case tested by MA-DRM in (Stoeve et al., 2009) were used for comparison purposes.

7.1.2 Problem Statement

Given a scenario of runway incursion as shown in Figure 7-1, the following information is considered known prior to the safety assessment.

1) Two Boeing 747-400 passenger aircraft: Aircraft Flying (AF) and Aircraft Crossing (AC), AF Pilots PF, AC Pilots PC, runway controller and infrastructures configured for the communication in runway control. See Figure 7-1a).
2) Aircraft configuration and performance specified by Table 7-1.
3) Runway configurations which are labeled in Figure 7-1b).
4) Communications specified by Figure 7-1b).

The objectives of implementing the CBSAF method is to 1) make quantitative safety assessments of systems with and without RIAS installed in runway operations and compare safety levels, with and without the stopbar and under two visibility conditions: a unrestricted visibility condition, Visibility Condition (VC) 1, and a restricted visibility condition, VC2.



a) Scenario Illustration b) System diagram (Stroeve et al., 2009)

Figure 7-1 Runway control

Table 7-1 Specification of Boeing 747-400 aircraft

Model	Boeing 747-400 Takeoff	Boeing 747-400 Taxi
Maximum takeoff weight	875000 lb (396,890 kg)	875000 lb (396,890 kg)
Engine Thrust /Engine	PW: 63,300 lbf (282kN) GE: 62,100 lbf(276 kN)	PW: 63,300 lbf (282kN) GE: 62,100 lbf(276 kN)
V_1: critical engine failure recognition speed	Flaps 20: 155 kts	--
V_r: rotate speed	Flaps 20: 171 kts	--
V_2: takeoff safety speed	Flaps 20: 180 kts	--
Takeoff distance	9900 ft	--
Wing span	64.4 m	64.4 m
Length	70.6 m	70.6 m
Taxing speed	--	Straight line: <20 kts

As shown in Figure 7-1, the safety of runway is monitored by runway controllers, pilots, and the automated system all of which are able to generate control commands and derive solutions to conflicts when needed.

For pilots, the only way to detect and dismiss any threats to aircraft safety is continuous visual surveillance, that is, by maintaining situational awareness. For air traffic controllers, the tower allows visual surveillance of the runway. For the RIAS, if installed, can provide two warnings: 1) stop bar violation warning, and 2) pending collision warning 15 seconds before the point of collision; and the radar display where movements of all aircraft are monitored and measured by radar system, also provides information about conflicts of two aircraft, if any. More details on the means of control are presented in Section 7.3. The following three sections demonstrate the implementation of the proposed CBSAF.

7.2 Stage I: Identify safety critical processes

7.2.1 Analysis on runway crossing process

See Figure 7-2. Define a coordinate system with the origin at the center of the intersection, $x-$axis along the take off runway centerline and $y-$axis along the crossing runway center line. Denote the positions of AF as (x_{AF}, y_{AF}), positions of AC to be (x_{AC}, y_{AC}).

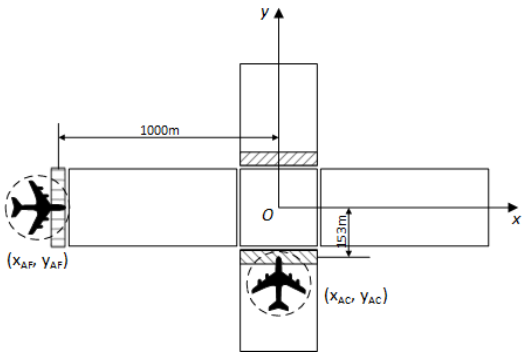


Figure 7-2 Runway Geometry and Coordinates

The initiating time instance is when either AF or AC starts moving towards the intersection. At this point, this moving aircraft (AF or AC) must have permission from the runway controller. The other aircraft (AC or AF) should not take off until the first aircraft has been cleared off. The process in question is between times when the second aircraft initiates takeoff or crossing procedure and when the runway has been cleared off.

Aircraft Flying (AF): Assume AF uniformly accelerates to takeoff at full power and maximum weight, then the acceleration is:

$$a_{max}^{+} = \frac{Max\ Power}{Max\ Weight} \approx 2.8\ m/s^2.$$

Where superscript "+" implies acceleration; for deceleration, "–" superscript will be used. Due to the drag as aircraft accelerates, assume that only 90% of the power was used for acceleration, hence,

$$a_{AF}^{+} \approx 90\% \times a_{max}^{+} \approx 2.5\ m/s^2.$$

From the initiation, it will then take a distance of $S_{v1}$to reach $v1$, the critical engine failure recognition speed, where

$$S_{v1} = \frac{1}{2}at^2 = \frac{1}{2} \times \frac{v1^2}{a_{AF}^+} = 1270m > 1000m.$$

Since this is larger than the distance between crossing runway and take runway starting line, it means AF can abort take off any time before reaching crossing runway. In fact, position of AF along the runway with respect to time can be estimated with

$$x_{AF} = \frac{1}{2}a_{AF}^+ t^2 - 1000$$

Should AF brakes with 70% of reserved full power at $x_{AF} = S_{AF}^*$, after S_b,

$$x_{AF} = S_{AF}^* + v_{S_{AF}^*}\left(t - t_{S_{AF}^*}\right) - \frac{1}{2}a_{AF}^-\left(t - t_{S_{AF}^*}\right)^2$$

Where, $v_{S_{AF}^*}$ is the speed of AF at S_{AF}^*, and a_{AF}^- is the decelerations at 50% reserved power:

$$a_{AF}^- = 50\%a_{max}^+ \approx 1.4\, m/s^2$$

Graphs with different S_{AF}^* values for v-t and S-t are plotted in Figure 7-3. These two figures can serve as look up tables for the different cases with different times when the imminent threat is detected.

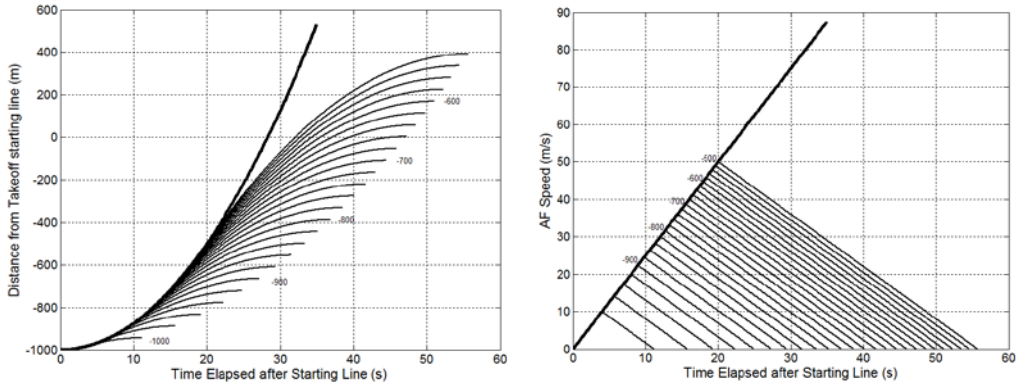


Figure 7-3 Distance and velocity over time for different braking *y*

Aircraft Crossing (AC) (Delta Virtual Airlines, 2009)**:** Similarly for AC, assume uniform acceleration at 90% of full power, until target taxi speed is met, after which, a constant taxi speed, will be maintained. The Delta pilot manual specifies the taxi speeds to be 20-30 kts for straight line (ICAO, 2007). In this case study $v_{AC}^{T} = 30\ kts$ ($15.4\ m/s$) will be used. Then, y_{AC} as a function of t is:

$$y_{AC} = \begin{cases} \frac{1}{2} a_{AC}^{+} t^2, t \le 6 \\ v_{AC}^{T} t, otherwise \end{cases}$$

Where, the acceleration

$$a_{AC}^{+} = 90\% a_{max}^{+} = 2.5 m/s^2$$

Assume AC could use 80% reserved power to come to a stop, then

$$a_{AC}^{-} = 80\% a_{max}^{+} = 2.2\ m/s^2$$

At any distance $y_{AC} = S_{AC}^{*}$, AC starts to stop, AC's position versus time is then,

$$y_{AC} = \begin{cases} S_{AC}^{*} + v_{S_{AC}^{*}} (t - t_{S_{AC}^{*}}) - \frac{1}{2} a_{AC}^{-} (t - t_{S_{AC}^{*}})^2, t \le 6 \\ S_{AC}^{*} + v_{AC}^{T} (t - t_{S_{AC}^{*}}) - \frac{1}{2} a_{AC}^{-} (t - t_{S_{AC}^{*}})^2, otherwise \end{cases}$$

The possible v_{AC} and y_{AC} over time t is visualized in Figure 7-4.

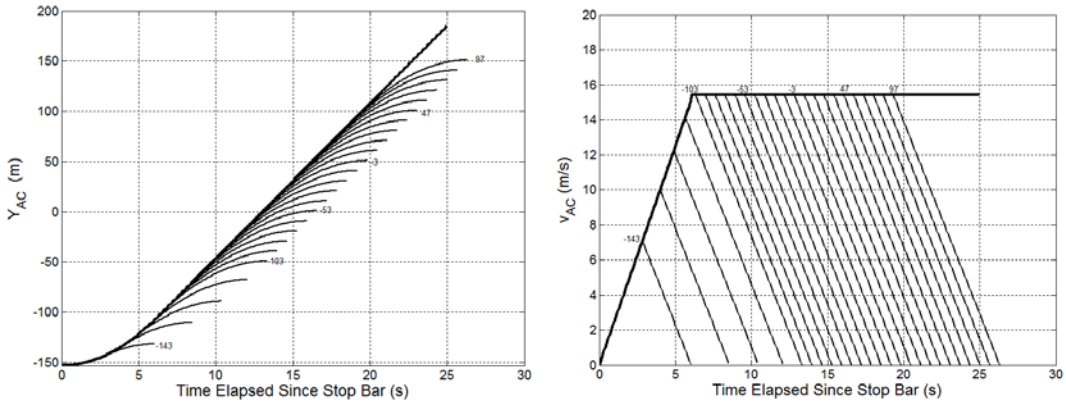


Figure 7-4 Distance and velocity over time for different braking *y*

The kinematics studies are necessary for state definitions. The level of specificity for aircraft kinematics analysis was determined by the needs to derive boundary values for defining the different states. State definition is the first step of Stage 1 in the theoretic method, in this case also needed for TSC calculations. It is also relevant to the PSC calculation in that the construction of control models needs information of the controlled safety critical process identified in stage 1 and depends on the state definitions.

CBSAF is a top down approach, which decomposes a system from control to the control functions to system elements. The lowest level of abstraction when applied to a given system is determined by the need to acquire the information required for calculating PSC and TSC. In the previous case, the TCAS research provided information about the state boundaries, therefore such details of kinematics analysis was not needed.

7.2.2 Define state space

Scope of the process to be studied are detailed as follows. First, only the realm of runway controller is considered; the ground controller's operations relevant to runway

will not be considered. Secondly, analysis is limited to AF and AC; possible collisions of ground vehicles are not accounted. Thirdly, AF follows nominal takeoff trajectory, and AC crossing trajectory, with negligible deviation. These analysis scope statements limit the analysis to only the two aircraft and not with nearby traffic or ground obstructions.

State space: Within the scope of analysis, define the state space of the system. Use the projected temporal separation t_S at the intersection of takeoff and crossing runway as the state variable, i.e. the time difference between the time stances AF and AC pass the intersections. At any given time, AF and AC are either on collision course ($t_S \approx 0$), or they are not ($t_S \gg 0$). In the cases AF and AC are not on collision course, they may either be safely separated, or that they are on the courses of near misses. See Figure 7-5.

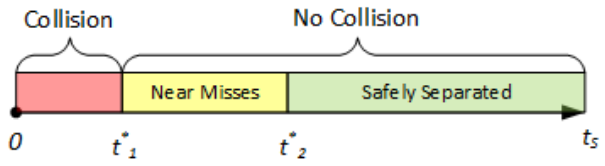


Figure 7-5 System States

Accounting for uncertainty in controlling the aircraft and the size of the aircraft, theoretical near misses, by $t_1^* = 2$ seconds for example may in reality likely result in a collision accident. Here $t_1^* = 2$ seconds, is arbitrary; it may vary case by case, but in general a very short time period at the order of seconds. Then, the following states may be defined:

X_1: $t_S = 0$, AF and AC collide;

X_2: $0 < t_S \leq t_1^*$, AF and AC on the collision course;

X_3: $t_1^* < t_S < t_2^*$, AF and AC on near miss courses;

X_4: $t_S \geq t_2^*$ AF and AC are safely separated.

t_1^* is the arbitrary critical value of temporal separation that borders X_2 and X_3 and t_2^* that borders X_3 and X_4. The state space is therefore $\{X_1, X_2, X_3, X_4\}$

Under uniform acceleration assumptions and using expected trajectories in Figure 7-3 and Figure 7-4, t_1^* is approximately 5 seconds and t_2^* approximately 34 seconds; both values are determined using worst case scenarios.

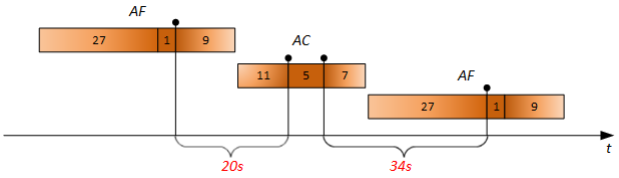


Figure 7-6 Temporal seperation at intersection

State transition space: Transverse all possible permutations, and manually check for viability of each transition. The acquired state transition space is shown in Table 7-3.

Table 7-2 State Transition Permutation

Initial State	$\rightarrow X_1$	$\rightarrow X_2$	$\rightarrow X_3$	$\rightarrow X_4$
X_1	--			
X_2		--	**	
X_3			--	
X_4				--

X_1 is the collision state, and thus a terminating state. Transitions in the first row of Table 7-2 therefore are not possible. X_4 is the state where two aircraft are safely separated,

by 34 seconds or more. Transitions from X_4 to X_2 will always first reach X_3, and transition to X_1 will pass X_3 and then X_2 before reach X_1. Therefore, although the transition physically is possible, based on the given state definition, they are not considered directly possible. Similarly, from X_3 to X_1, direct transition is not possible, since X_2 need to first be reached. For the transition from X_2 to X_4, first X_3 has to be reached. The impossible transitions are colored grey in Table 7-2.

Safety critical process: First, have all possible state transitions listed in the first column of Table 7-3. Safety critical processes are controlled transitions which without control will directly evolve into an accident. The control-less transition $X_2 \rightarrow X_1$ is the transition to accident, and with control, X_2 may transit to X_3. Therefore the safety critical process is $X_2 \rightarrow X_3$, also marked "**" in Table 7-2.

Table 7-3 State transitions of runway incursion

State transitions	Type
$X_2 \rightarrow X_1$	unsafe process
$\boldsymbol{X_2 \rightarrow X_3}$	With control; safety critical process
$X_3 \rightarrow X_2$	Performance deviation; proceeding unsafe process
$X_3 \rightarrow X_4$	With control; critical process
$X_4 \rightarrow X_3$	Performance deviation; proceeding unsafe process

Table 7-3 also shows the identified safety critical process: $X_2 \rightarrow X_3$, the the transition from X_2 (AF and AC on collision course) to X_3 (AF and AC near miss by $t_s \in (t_1^*, t_2^*)$. Without any control actions, X_2 will evolve into the accident state X_1.

Compare the transition space with the unsafe scenarios identified in (Blom et al., 2006):

- Scenario I: Aircraft erroneously in take-off and crossing aircraft on runway
- Scenario II: Aircraft erroneously crossing and other aircraft in take-off
- Scenario III: Aircraft taking off and runway unexpectedly occupied;
- Scenario IV: Aircraft crossing the runway and runway unexpectedly occupied by aircraft
- Scenario V: Aircraft crossing and vehicle on runway;
- Scenario VI: Collision between aircraft sliding off the runway and aircraft near crossing;
- Scenario VII: Aircraft taking off and vehicle crossing;
- Scenario VIII: Jet blast from one aircraft to another; and
- Scenario IX: Conflict between aircraft overrunning/climbing out low and aircraft using a nearby taxiway.

Under the analysis scope of applying CBSAF, Scenarios V, VI, VII, VIII and IX are excluded. The scenarios defined with MA-DRM method enumerates violations to standard procedures, or hazards in the system that may pose threats to the runway operation safety. The completeness of all possible scenarios relies heavily on the analyst's expertise and thus can be arbitrary. The CBSAF assures that the state definitions are complete, but it may cover several scenarios where the control options are different. X_3 the near miss state, for example, can be scenario I AF takes off after AC actively crossing the runway, or Scenario II, AC starts crossing before AF has cleared off the runway. The resolution for the two scenarios may be different. In Scenario I, since AC has already left the stopbar, the stopbar lights could not be used to stop AC. In scenario II, AF is in takeoff, AC is still at the stopbar and therefore, the stopbar lights can be turned on to warn PC.

On the other hand, the scenarios listed are not necessarily unsafe events. For example, in Scenario I, due to the configurations in question, if no control is applied, AC will separate AF by at least 11 seconds, and therefore will remain X_3. Similarly, Scenario II may be classified as X_2, X_3 or X_4.

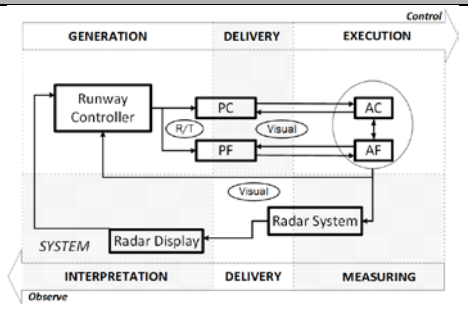
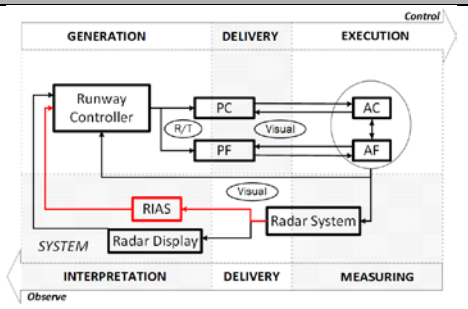
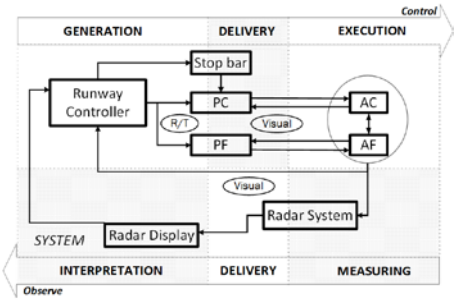
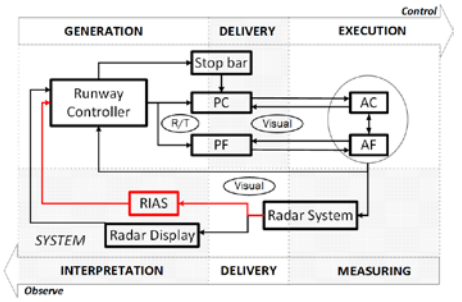
In CBSAF, the choice of state variable is not straightforward. In this case study, instead of using t_s the temporal separation at the intersection, an alternative way is to use the two aircraft's physical position and velocity (whether $v = 0$). The enumeration however will be within a large set of possible states and state transitions.

7.3 Stage II: Control Models

Recall the guideline of system modeling in Figure 5-7. Use CBSAF to identify the system's control models over different processes. Based on the six control functions and the feedback control loop structure, the corresponding hardware, software (interactions, procedures etc.) that enable these functions are identified and the control structure developed.

For the identified safety critical transition p_{cr}: $X_2 \rightarrow X_3$, there may be different scenarios based on two configuration differences: 1) whether AC has moved passed the stopbar, and 2) whether RIAS is installed in the system. This gives four systems with different control mechanims: 1) with stop bar and RIAS, 2) with stop bar and no RIAS, 3) without stop bar and with RIAS, and 4) without stop bar or RIAS. The developed control models for the four systems are illustrated in Table 7-4.

Table 7-4 Control Models of $p_{cr}: X_2 \rightarrow X_3$

	Without RIAS	With RIAS
Without Stopbar		
With Stopbar		

Means of control: As shown in the control models, there are three decision makers: runway controller(s), Pilot Flying (PF), and Pilot crossing (PC). The redundant components also yield several means of controls enabled by connections, which are listed in Table 7-5.

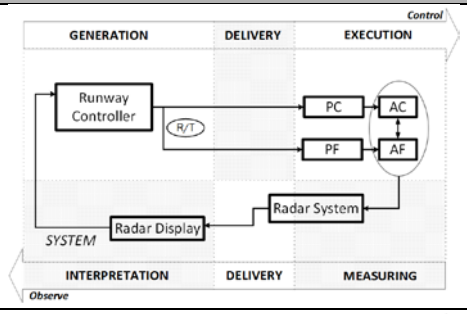
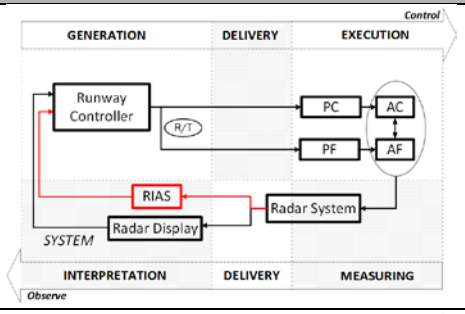
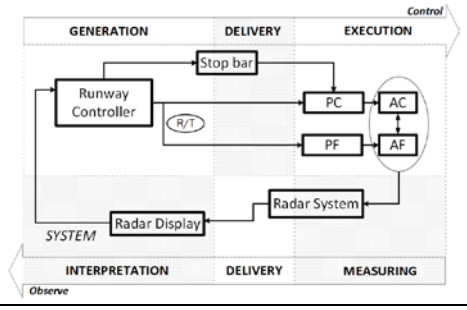
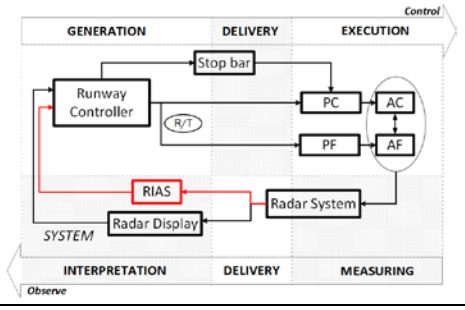
As shown in the control models, the three control routes are not independent of each other. For example, one means of control is through PC see and avoid, and a second is through ATC controlling PC to avert collision. PC is shared between the means of control, without which both means will be lost. Although all of them are able to control at least one of the safety critical processes above, due to the interdependency, control capacity must be evaluated as one control system.

Table 7-5 Means of Control

Means	Sub. Cat.	Means of Control
ATC	visual	ATC Visual → R/T → PF → AF ATC Visual → R/T → PC → AC ATC visual → Stopbar → PC → AC
	visual	ATC Radar Display → R/T → PF → AF ATC Radar Display → R/T → PC → AC ATC Radar Display → Stopbar → PC → AC
	RIAS I	ATC RIAS stop bar violation alert→ R/T → PF → AF ATC RIAS stop bar violation alert → R/T → PC → AC
	RIAS II	ATC RIAS conflict alert → R/T → PF → AF ATC RIAS conflict alert → R/T → PC → AC
`PF	Visual	PF Visual → R/T → PF → AF
PC	Visual	PC Visual → R/T → PC → AC

Visibility conditions: The MA-DRM considers two visibility conditions: 1) unrestricted: both pilots and ATCo can visually observe the traffic situation; and 2) visibility range between 400 m and 1500 m: ATCo cannot visually monitor traffic and the pilots are not always able to see the other aircraft during the first part of the takeoff or crossing. Condition 2 indicates foggy weather around runways. In this study, assume with Condition 2, controls rely solely on the radar systems, and hence visual surveillance are unavailable, even though the given visibility range covers a slight chance where PF and PC can see each other. Control models with this assumption are shown in Table 7-6.

Table 7-6 Control Models of $X_2 \rightarrow X_3$ for Visibility Condition 2

	Without RIAS	With RIAS
Without Stopbar		
With Stopbar		

As shown in Table 7-6, compared to unrestricted visibility conditions, with restricted visibilities, feedback from ATCo visual monitoring is no longer available. Additionally, PF and PC are unable to detect conflicts, and act as actuators only.

7.4 Stage III: Evaluate System Control capacity

7.4.1 Probabilistic System Control capacity

Along each possible control loop, the events to achieve the six functions (i.e. sensing, delivery and interpretation of process information and generation, delivery and execution of control commands) follow a sequential order. Event Tree Analysis is therefore suited to obtain permutations for all possible chains of events that lead to preservation of at least one means of control, *PSC*, or loss of all means of control *1-PSC*. The individual events, or pivotal events, are identified via hazard identification based on the control models. The hazard list derived from the hazard identification is then used

for event tree construction. The quantification of PSC also requires the probabilities of occurrences of each hazard event. As stated in Chapter 4, these values of probabilities of occurrence of the hazards are the assumed performance deviations and applied across the different control systems for comparison purposes only. The outcome of a system's PSC by itself from this approach is not absolute risk.

7.4.1.1 Hazard Identification and specification

Hazard identification is conducted given the controlled process $p_{cr}: X_2 \rightarrow X_3$ and its control models. Table 7-7 shows the hazard event set identified and used for PSC evaluations. For comparison, the hazard list from (Stroeve et al., 2013) on similar runway incursion case study is shown in Table 7-8.

Table 7-7 Hazard list derived by CBSAF

ID	Control Function	Type	Event Description	P Used
H1	Sensing	S	ATCo fails to detect conflict visually	0.5
H2	Sensing	S	PC fails to detect conflict visually	0.5
H3	Sensing	S	PF fails to detect conflict visually	0.5
H4	Delivery	H	Radar system fails	0.01
H5	Delivery	H	RIAS fails to warn ATC stop bar violation	0.01
H6	Delivery	H	RIAS fails to warn ATC collision at 15s to accident	0.01
H7	Interpretation	S	ATCo fails to detect conflict on radar display	0.3
H8	Generation	S	ATCo fails to generate solution	0.1
H9	Delivery	H	R/T fails	0.01
H10	Delivery	S	PC fails to conform to stop bar warning	0.05
H11	Delivery	S	ATCo fails to communicate effectively with PC	0.2
H12	Delivery	S	ATCo fails to communicate effectively with PF	0.2
H13	Execution	S	AC fails to effectively resolve the conflict	0.1
H14	Execution	S	AF fails to effectively resolve the conflict	0.1

Table 7-8 Hazard list from (Stroeve et al., 2013)

ID	Control Fun(s)	Event Description	P min	P max
Q1	Environment	No aircraft in take off	0.75	0.75
Q2	Sensing –exec.	Pilots recognize and resolve conflict at early stage	0.5	0.7
Q3	Sensing	Controllers recognize conflict at early stage	0.1	0.2
Q4	Sensing	Alert System warns controller at early stage	0.95	0.99
Q5	Ctrl Del.-exec.	Communication leads to resolution at early stage	0.8	0.9
Q6	Sensing – exec.	Pilots recognize and resolve conflict at medium stage	0.9	0.99
Q7	Sensing –interp.	Controllers recognize conflict at medium stage	0.2	0.4
Q8	Sensing	Alert system warns controller at medium stage	0.9	0.99
Q9	Ctrl del. – exec.	Communication lead to resolution at medium stage	0.6	0.8
Q10	Sensing – exec.	Pilot recognize and resolve conflict at late stage	0.9	0.99
Q11	Sensing – interp.	Controllers recognize conflict at early late stage	0.5	0.75
Q12	Ctrl del. – exec.	Communication leads to resolution at late stage	0.4	0.6

Comparing Table 7-7 with Table 7-8, the general hazard types are the same: ATCo detects conflict, alert system warns ATCo, ATCo command solution through R/T, and pilots see and avoid. Discrepancies also exist however. Compared to CBSAF, the list given by (Stroeve et al., 2013) contains more combinatory hazards and is less detailed. For instance, the counterpart of Q2 in CBSAF may be the combination of H2 and H13.

Another visible difference is the repetition of early, medium and late stages of the same hazards used in MA-DRM. Since the CBSAF assumes only one control action is taken in the process to be analyzed, the “late stage”, the earlier and medium stage events of the same natures are beyond one control cycle and not taken into account by CBSAF.

7.4.1.2 Event Tree Analysis

Use ETA to derive PSC. Figure 7-7 is an example of the event tree developed. Due to the oversize of other event tree diagrams, they are not shown.

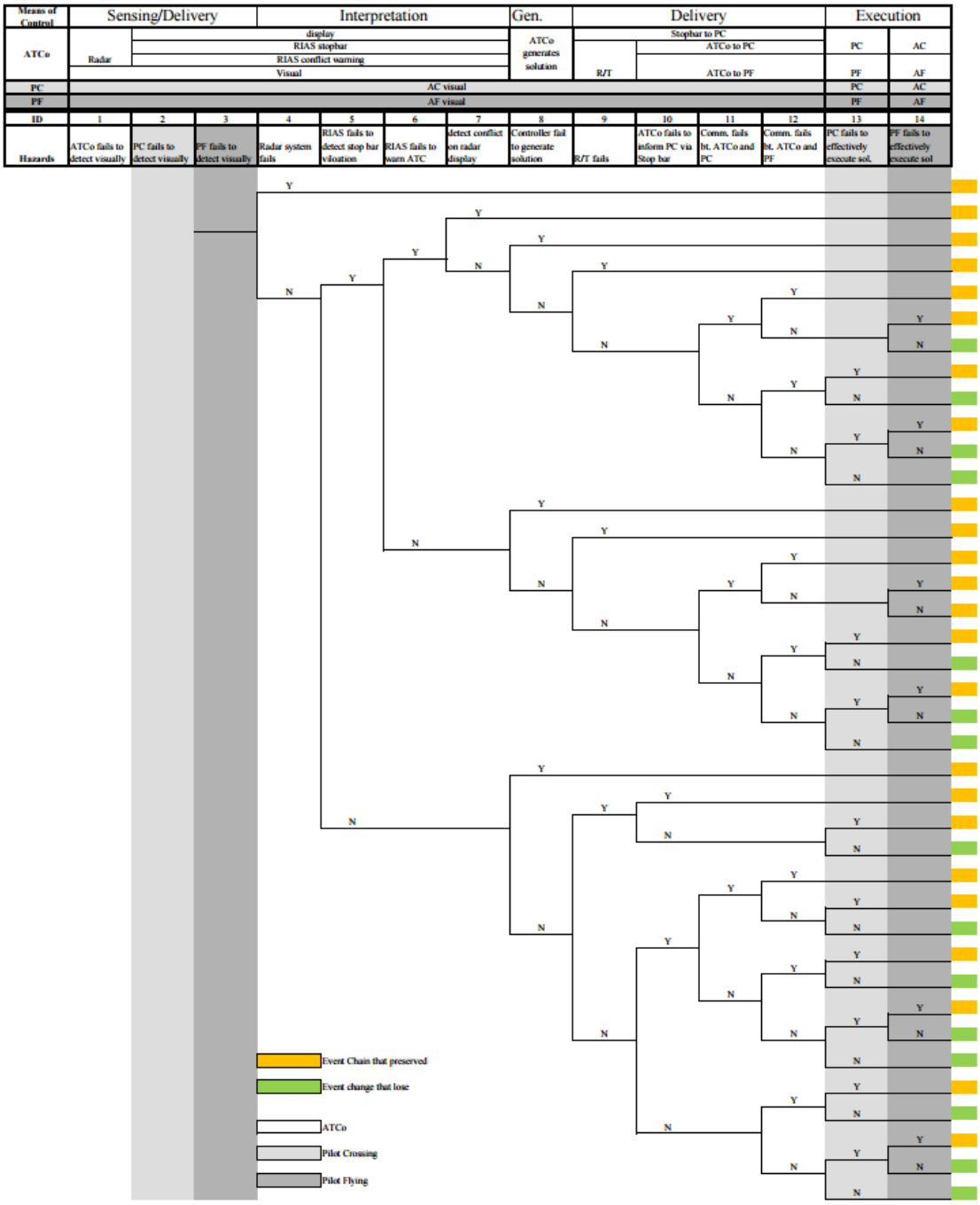


Figure 7-7 Event Tree Diagram for Visibility Condition 2

PSC is the sum of probabilities of the event chains represented by braches that lead to preservation of at least one means of control. Repeat the process for the total of 8 control scenarios. Assuming along the timeline, the events are independent and probabilities of occurrence are as shown in the Figure 7-4, the resultant PSC are summarized in Table 7-9 and Figure 7-8.

Table 7-9 PSC for all 8 control scenarios

		Without RIAS	With RIAS
Visibility Condition 1	Without stop bar	0.9074	0.9085
	With stop bar	0.9214	0.9250
Visibility Condition 2	Without stop bar	0.5690	0.8129
	With stop bar	0.5691	0.8623

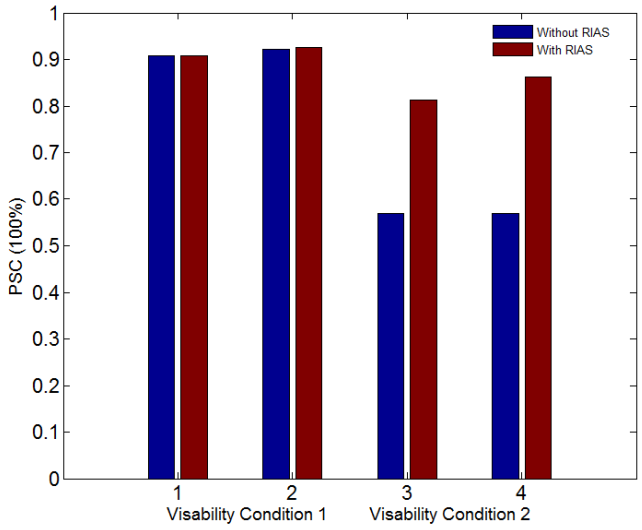


Figure 7-8 PSC for with and without RIAS

7.4.2 Temporal System Control capacity

7.4.2.1 Time available t_A

The safety critical process presumes that without control AF and AC will enter the intersection at almost same time ($t_S \approx 0$).

Process $x_1 \rightarrow x_2$ requires either aircraft to come to a stop. For this to be possible, The latest time for AF to come to a full stop according to Figure 7-3 is if AF initiates the stop at $x_{AF} = -740m$. And for AC to be able to stop before the runway, AC needs to act at or before $y_{AC} = -53m$. The distance between the two at this point is

$$d_{FC} = \sqrt{x_{AF}^2 + y_{AC}^2} \approx 731.9m$$

Recall Figure 7-3 and Figure 7-4, it takes AF about 1 second to cross the intersection, and it takes AC about 4 seconds to cross the intersection. This causes AC to start crossing about 11 seconds after AF has initiated take off. The control is to interrupt this process and stop two aircraft. Ideally both aircraft should be stopped. But if at least one aircraft manage to stop, the accident will be averted.

If the aircraft are on a collision course, at least one aircraft has taken action to stop, then $t_A = 17s$, all three means of control, i.e. through ATC, through PF and through PC. If the visibility is lower than 731.9m, then only ATC can be used.

7.4.2.2 Time required t_R

Recall Equation 5.1, and apply the procedure recommended by Figure 5-11. Each time components for t_R is summarized in Table 7-10.

Table 7-10 t_R calculation summarized

Means	ATC to AF or AC			PF-AF		AC-PC	
Gen.	Runway controller		1-2s	See and avoid	0.5- 1s	See and avoid	0.5 -1s
Del.	T/P to PF or PC	Stop bar	1-5s				
Exe.	PF maneuver	PC maneuver	2-5s	Maneuver AF	2-5s	Maneuver AC	2-5s
Sen.	Visual	Radar	0.01–3s	Visually Confirm	0.5 -1s	Visually Confirm	0.5-1s
Del.		Radar to tower					
Interp.		RIAS	N/1-3s				
t_R	5.01 – N s			3-7s		3-7s	

Note: N is a positive large value. This value is used to account for the possibility that the state is detected very early.

7.4.2.3 TSC

In the presence of RIAS, the results of TSC for the different system configurations to control the safety critical process $x_1 \rightarrow x_2$ are listed in Figure 7-2.

Table 7-11 TSC for all Safety Critical Processes

	RIAS	t_A	t_R	TSC
$x_1 \rightarrow x_2$	Y	17s	(3, 17)s	(0, 14)s
	N	17s	(3, N)s	(N-17, 14)s

Note: N is a positive large number and N>17s.

7.5 Result analysis

7.5.1 Sensitivity Test

Since the probabilities of occurrence for the element hazard event are used to derive PSC of the four system configurations, whether the assumptions on such performance

deviation will affect the comparison qualitatively should be examined. In this section, the assumptions on hazard event probabilities are tested. The ranges of probabilities used for the events are listed in Table 7-12.

Table 7-12 Table 7-12 Hazard list and probability ranges in sensitivity tests

ID	Control Function	Type	Event Description	P_0	P_{min}	P_{max}
H1	Sensing	S	ATCo fails to detect conflict visually	0.4	0.3	0.7
H2	Sensing	S	PC fails to detect conflict visually	0.4	0.3	0.7
H3	Sensing	S	PF fails to detect conflict visually	0.4	0.3	0.7
H4	Delivery	H	Radar system fails	0.01	0.01	0.1
H5	Delivery	H	RIAS fails to warn ATC stop bar violation	0.01	0.01	0.1
H6	Delivery	H	RIAS fails to warn ATC collision at 15s to accident	0.01	0.01	0.1
H7	Interpretation	S	ATCo fails to detect conflict on radar display	0.3	0.1	0.3
H8	Generation	S	ATCo fails to generate solution	0.1	0.1	0.3
H9	Delivery	H	R/T fails	0.01	0.01	0.1
H10	Delivery	S	PC fails to conform to stop bar warning	0.05	0.05	0.1
H11	Delivery	S	ATCo fails to communicate effectively with PC	0.2	0.01	0.3
H12	Delivery	S	ATCo fails to communicate effectively with PF	0.2	0.01	0.3
H13	Execution	S	AC fails to effectively resolve the conflict	0.1	0.01	0.3
H14	Execution	S	AF fails to effectively resolve the conflict	0.1	0.01	0.3

For each test, a probability value for each hazard event is taken from the given range under uniform distribution assumptions. A total of 5000 random combinations were used to find the variations of PSC. The PSC distribution for Visibility Conditions 1 is shown in Figure 7-9 and the PSC distribution for the same combinations are shown in

Figure 7-10. The distribution of the differences for with and without RIAS for four groups of scenarios are shown in Figure 7-11: 1) VC1 and without stopbar, 2) VC1 and with stopbar, 3) VC2 and without stopbar and 4) VC2 and with stopbar.

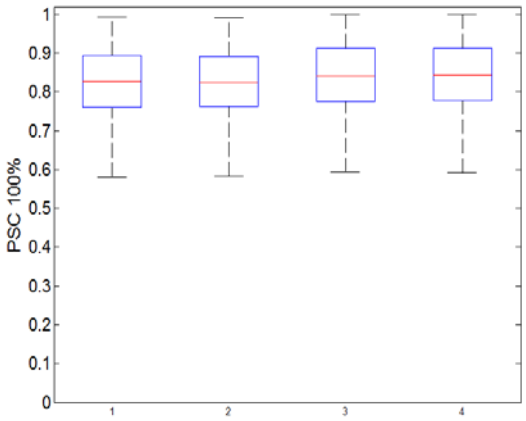


Figure 7-9 Visibility Condition 1

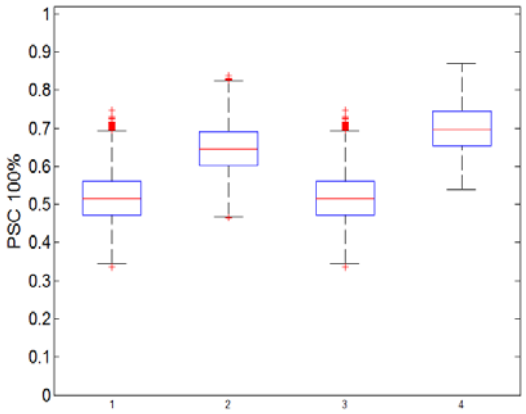


Figure 7-10 Visibility Condition 2

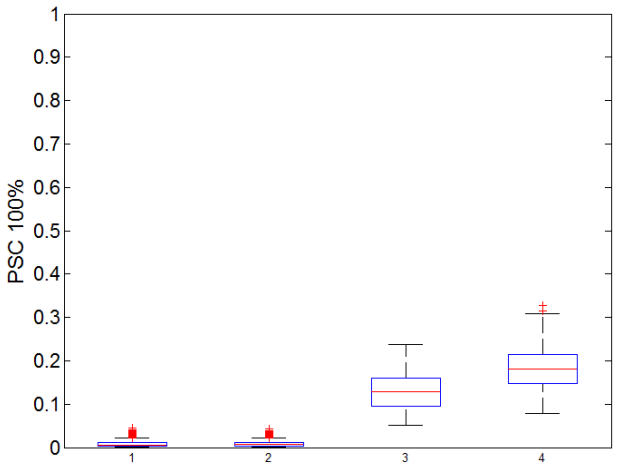


Figure 7-11 Difference for With and Without RIAS

As is seen from Figure 7-10, for the range of probabilities selected, the PSC values are between 0.55 to 1 with the average of about 0.8 for all four cases: with/without RIAS and with/without stopbar. In comparison, for Visibility Condition 2, Figure 7-11 shows ranges of PSC for with RIAS are about 0.33 – 0.7, and without RIAS is about 0.45 – 0.85.

The difference for with and without RIAS are more clearly illustrated in Figure 7-11, where in Visibility Condition 1, the difference is between 0 to 0.04, whereas in visibility condition, the difference is between 0.06 to 0.3.

For the ranges of probabilities of the elementary hazard events assumed in Table 7-12, the sensitivity test results therefore support the qualitative comparison results: the safety benefits using the PSC measure indicates negligible increase for VC 1 and more significant for VC2.

7.5.2 Comparison with MA-DRM

PSC: The results obtained by CBSAF indicate that no substantial improvement of control capacity is observed with RIAS applied under VC 1 and more significant with VC 2. Under VC 1 and for without stopbar, the PSC values for the configurations from without to with RIAS has a small increase of 0.1%. In comparison, under VC 1 and for without RIAS, PSC of the system from without to with stopbar has an increase of 1.5%, and for with RIAS, PSC from without to with stopbar an increase of 1.8%.

Comparing the different visibility conditions, under VC 1, unrestricted visibility, the average increase for the different combinations of performance deviations assumptions, the PSC increase is at the order of 0.001. Whereas, under VC 2, restricted visibilities, the PSC on average increase by 0.1 to 0.2.

Since the controlled process $p_{cr}: X_2 \rightarrow X_3$ is safety critical, the PSC evaluations are expected to correlate to the system's safety performance. The comparisons indicate that under VC 1, the introduction of RIAS does not substantially increase safety, and under VC 2, the introduction of RIAS has a more significant impact.

The cases of stopbar as additional quantitative reference indicate that between stopbar and RIAS under VC 1, stopbar has a higher influence on the PSC and the safety performance. With respect to system design, this is an implication that the addition of stopbar in VC 1 is a more effective safety assurance strategy.

To interpret the small increase of control capacity, PSC, RIAS is an additional means to the many other means of detecting the pending collision. Under VC2, since many existing visual means of detection are hindered or disabled due to the reduced visibility conditions, the introduction of RIAS becomes more significant.

The same conclusions are reached in the case study using the MA-DRM method as is seen in Figure 7-12, the difference for with and without RIAS is more significant for VC2 compared to VC1 based on the risk outcomes. The CBSAF in comparison is computationally less costly; it also adds a new control interpretation to the outcomes from the analysis.

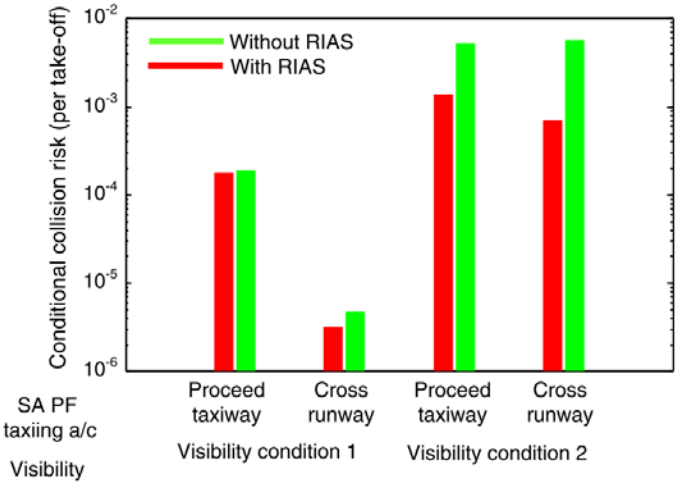


Figure 7-12 MA-DRM method results (Stroeve et al., 2009)

TSC: Since only the CBSAF has temporal measures, TSC is not used in the comparison but as additional references. For TSC, if the worst case is considered, that none of the human components is aware of the pending collision, then RIAS provides increased assurance that the controller has at least 15 seconds to respond. Namely, time available $t_A \geq 15s$. Comparing time available time required for the system to respond and react, TSC has to be greater than 0 to accommodate frequent delays in operations. In VC2, TSC is not guaranteed to be greater than 0, which makes the introduction of RIAS necessary to assure a positive TSC.

Considering other means of increasing time available for early detection and in the resolving runway incursion threats, the higher the TSC, the safer the system. Therefore the TSC is an important references that gives a different aspect on system control capacity and safety consideration to the evaluating the safety benefits of introducing the RIAS system.

CHAPTER 8. CONCLUSION AND RECOMMENDATION

8.1 Conclusion

8.1.1 Theoretical foundations

Based on the view that “system safety is a control problem”, the concept “control capacity” can be correlated to a system’s safety performance, through the control of safety critical processes. Safety critical processes are processes whose failure to obtain its objective will result in safety consequences, e.g. an accident. The control capacity of a system in the control of safety critical processes is an indicator of its safety performances. This theoretical elicitation is based on the research literature on system safety and system control, and used as the foundations to the use of “control capacity” as a system safety performance measure.

8.1.2 Control capacity and metrics

From a safety perspective, system control capacity is defined as the extent for a system to withstand performance deviation in obtaining its control objectives. While control systems as well as performance deviations are multi-faceted, this definition is applicable to any control systems, and any type of performance deviations.

In order to quantify the performance deviation and therefore the extent to withstand the performance deviation however, the research focused on two types of performance

deviations: 1) failure or faults and 2) delays. This approach is demonstrative for other types of performance deviation and may serve as a reference for an integrated approach to performance deviation examination. The two metrics to quantify system control capacity proposed are “Probabilistic System Control Capacity” and “Temporal System Control Capacity”, to address the two types of performance deviation respectively. The justification for the selecting the particular two aspects are that in the interactions between the controller and the controlled process, PSC has an emphasis on the controller side and TSC has one on the controlled process.

8.1.3 CBSAF

To adapt the theoretic measure “control capacity” and its metrics PSC and TSC to the quantitative safety assessment of the ATC systems, a Control-capacity Based Safety Assessment Framework (CBSAF) is needed and assembled. In addition to the two metrics, the research elicited a three stage procedural method that consist of principles, guidelines and procedures for setting up a QSA for an ATC system: I) identify safety critical processes, II) develop control models, and III) evaluate PSC and TSC. The procedure is tailored toward QSA of ATC systems, but can be adaptable to QSA of other types of safety critical systems, as the principles and guidelines are developed to be as general as possible.

8.2 Case studies

8.2.1.1 Utilities and Potentials of CBSAF

Two case studies collision avoidance and runway incursion applying CBSAF demonstrated utilities and potential of the CBSAF method. In the collision avoidance case study, three hypothetical configurations are set up for comparing the effect of activating two means of control simultaneously. The configurations are hypothetical because they are not the standard procedures where the two means of control are regulated to activate at different times and circumstances. Comparing Configuration II of air traffic controller only Configuration III with both air traffic controller and the automated system TCAS, there was an observed decrease in control capacity. And since collision avoidance is a safety critical process, this decrease in control capacity is then an indicator that the system's safety performance is compromised with the addition of air traffic controller: from Configuration II to Configuration III. This is aligned with the observation of the Uberlingen Mid-air collision accident.

Stability of the assumptions on performances deviation in terms of probabilities of occurrences of the elementary hazardous events are tested in two sensitivity test, one with varying intervals of uncertainty and the other with uniform intervals. The sensitivity tests supported the qualitative conclusion drawn from the comparison using the original assumed performance deviation assumptions. The comparisons among the three systems are robust to the variations in assumptions of the probability values within the

reasonable range. The second sensitivity test shows that for a set of unlikely combinations of probabilities, the PSC may lead to the opposite conclusions.

In the second case study on a specific runway incursion scenario, the CBSAF was compared to an existing QSA method MA-DRM. The same settings from MA-DRM study case are adopted to compare ATC systems with and without RIAS and under unrestricted and restricted visibility conditions. Following the CBSAF methodology, the results are able to reach similar conclusions as using the MA-DRM with less computational costs, that under unrestricted visibility conditions the safety benefit of RIAS is negligible and under restricted visibility conditions more observable. Similar to Case Study I, the quantity assumptions on the probabilities of occurrence of identified failure/faults hazard events are tested for stability. The results show that the comparison conclusions hold for a wide range of quantities of the probabilities.

The CBSAF also provides a new control perspective to the interpretation of this finding. In the unrestricted conditions, RIAS is added to existing four means of the observe part (sensing/delivery/interpretation), which is first half of the control loop. Therefore the impact of the redundant mean of observing is less significant than if it was the one of fewer, e.g. only means of observing. Under the restricted visibility condition, the visual means of observing are hindered or disabled, which then makes the means of observing through RIAS more impactful to the first half of the control loop.

8.2.1.2 Uncovered issues of CBSAF

Constricted assumptions: Both PSC and TSC are measured under the assumption that only one control cycle is used in the control of a safety critical process. The assumption is overly simple to account for situations when one control means is able to attempt for a second control cycle when all control means failed in the first control cycle. For example, in (Stroeve, 2009), early, medium and late detections were taken into account. In the case studies, due to constraints on resources and limited access to data, the assumptions to acquire time components of TSC and for definition states require further examinations. The TSC results may be treated as demonstrative; they are not recommended to use as reference for real systems.

In PSC evaluations, the assumptions on the probability values lead to restricted applicability of CBSAF to comparisons between similar control systems over the same safety critical processes. The definition of PSC is extendable to when more accurate probabilities are available, e.g. through data, in which case, PSC can be used for comparison between significantly different system and for even calibration of target levels of safety.

Limited scalability: First step in Stage 1 of CBSAF starts from selecting state variables. Both case studies used 1 temporal variables that result in a small set of states. In other cases where the state variables are large in quantities, the possible combinations of state variables that characterize the system expand quickly. In other words, the state definition has limited scalability to complex scenarios that require non-trivial number of states.

A similar problem exists in the use of event tree analysis. As discussed in Section 5.4, the ETA technique was selected for its simplicity and intuitiveness; equivalent quantitative risk assessment approaches, e.g. FTA, can be used for this step as well. The scalability issue is inherent to ETA. As the number of hazard increases, the ETA size will grow rapidly as well. The probabilities if very small, will incur roundabout errors using computational tools, e.g. a PC.

Subjectivity: Safety assessments are intrinsically subjective and rely heavily on experts and assumptions about the system. The CBSAF also has a number of subjectivity concerns uncovered in the case studies, as are listed:

1) The definition of states requires arbitrary criteria and is subject to individual variations.
2) The construction of control models requires manual identification of system elements and mapping of the elements to the control functions.
3) The hazard identification relies on levels of expertise and spectra of experiences of the analysts and could also have individual variations
4) The construction of event tree diagram will apply subjective terminating criteria, i.e. whether a combination of hazard/failure event outcomes will determined to lead to success or failure of control.
5) The probabilities of hazard outcomes are assumptions determined by the analysts and therefore will also contain individual difference, due to the varying levels of expertise and experiences with the system.

These subjectivity issues are common in QSA and use of the ETA technique. Preliminary attempts including sensitivity tests are adopted to investigate the impact of the subjectivity factors. It is recommended further and more thorough studies be conducted on subjectivity before introducing the CBSAF to implementation phases.

8.3 Future work

8.3.1 Additional Control Capacity Measures

Control capacity in this research was defined as a tolerance of performance deviations in the control of safety critical processes. Performance deviations are of many forms and scales. Other than failure, faults and delay, performance deviation may be of other kind for example, information integrity through the control loop and lack of coordination. It is recommended that the many facets of system control be explored and other control capacity measures be proposed and studied.

8.3.2 Alternative Control Capacity measures

Control capacity is a general attribute common to all control systems, which is a system's capacity to control. The definition used in this research takes the perspectives of safety and performance deviation and measure control capacity by different types of performance deviation.

The more direct approach can take two directions. First, the different performance deviations are integrated into one measure. In the STAMP based Process and Analysis (STPA), a taxonomy of control faults is given, which covers a range of performance deviations. The integrated approach may use this reference to strive for a measure that accommodates different types of performance deviations.

The second direction for directly measure control capacity is to investigate the factors that affect a system's capacity to control. In Section 3.1.1, two illustrative examples are given to show that this system control capacity vary with different systems/controllers, and the different processes to be controlled. A number of factors are identified and

listed that contribute to the different control capacities including redundancy and directness. This second recommended direction can further this effort and explore the control capacity measure directly from the contributing factors to the variations of system control capacity.

8.3.3 Automated algorithms

Another recommended improvement of the method is to introduce automated process to the estimation of PSC, which requires construction of large size event trees. Although the process is not entirely mechanical and cannot be fully automated, there is a level of automation that can speed up the process and reduce errors induced in the manual development of event tree diagrams. For example, the permutation of hazard events, after being sorted in chronological order, is suited for the capacity of a computer. A user interface can be developed to prompt the sequence of events to analysts and request for approval. Additionally, some simple rules can be used for repetitive subscenarios, with simple rules. For examples, if there is two sensor systems, I and II, for the case when system I fail and when system II fail, the rest of the system will have the same delivery-interpretation-generation-delivery-execution (sub)scenario and thus the same probability of failure.

Another benefit for accommodating larger size problem is the increased capacity of more hazard events. In problems with more complex procedures and components, this may be critical for the CBSAF to be feasible.

8.3.4 Complex control models

Large scale, complex and socio-technical systems often have many levels of control structures. The hierarchy and interdependence requires comprehensive and systemic examinations. The study of controls of controls may very well be a topic itself. It is may be connected to management, when the main components are humans.

Control is of scientific subject. The science of control is universal and applicable to an extremely broad range of processes, systems and even organizations. The generalized theories of control mechanism therefore are possible as much as useful. The interested readers are encouraged to further explore and formulate general theories of control, in the context of general systems and applicable to safety critical socio-technical systems.

8.3.5 Validation and Verification

Both case studies in this research are for demonstrative purposes with limited verification objectives. The CBSAF framework is subject to thorough and rigorous validation and verification processes before it reaches practical and implemental phases. For example, the scalability and subjectivity issues discussed in the conclusion session needs further study on their impacts. Alternative and more advanced quantitative risk assessment techniques can also be adapted and tested in this framework.

The assumptions used for deriving the CBSAF and for the case studies require further examinations, both for their practicalities and whether they can be loosened to accommodate expanded use of the CBSAF, e.g. for comparison between system that are significantly different and even across domains such as between ground transportation and air transportation

REFERENCES

REFERENCES

Airplanes, C. 2015. Statistical Summary of Commercial Jet Airplane Accidents.

Amalberti, R. 2001. The paradoxes of almost totally safe transportation systems. Safety Science. 37: 109-126.

Andrews, J. 1998. Fault Tree Analysis. Proceedings of the 16th International Safety Conference.

Andrews, J. D. and Dunnett, S. J. 2000. Event-tree analysis using binary decision diagrams. Reliability, IEEE Transactions on. 49: 230-238.

Australia, S. W. 2012. Guide for Major Hazard Facilities - Safety Assessment.

Bass, T., and Robichaux, R. 2001. Defense-in-depth revisited: qualitative risk analysis methodology for complex network-centric operations. In Military Communications Conference. MILCOM 2001. Communications for Network-Centric Operations: Creating the Information Force. IEEE. Vol. 1, pp. 64-70.

Bahill, A. T. 2012. Diogenes, a process for finding unintended consequences. Systems Engineering. 15: 287-306.

Baraldi, P. and Zio, E. 2008. A combined Monte Carlo and possibilistic approach to uncertainty propagation in event tree analysis. Risk Analysis, 28**,** 1309-1326.

Becker, J. C., and Flick, G. 1997. A practical approach to failure mode, effects and criticality analysis (FMECA) for computing systems.

Bertalanffy, L. V. 1968. General system theory: Foundations, development, applications

Blom, H., Bakker, G., Blanker, P., Daams, J., Everdij, M. and Klompstra, M. Accident risk assessment for advanced air traffic management.

Blom, H. A., Stroeve, S. H. and De jong, H. H. 2006. Safety risk assessment by Monte Carlo simulation of complex safety critical operations. Developments in Risk-based Approaches to Safety. Springer.

Boulding, K. E. 1956. General systems theory-the skeleton of science. Management science, 2(3), 197-208..

Brooker, P. 2002. Future air traffic management: quantitative en route safety assessment .1. Review of present methods. Journal of Navigation. 55**:** 197-211.

Brooker, P. 2002. Future air traffic management: quantitative en route safety assessment Part 2-New approaches. Journal of Navigation. 55**:** 363-379.

Brooker, P. 2004. Consistent and up-to-date aviation safety targets. Aeronautical Journal. 108**:** 345-356.

Brooker, P. 2007. Are there good air traffic management safety indicators for very safe systems? Journal of Navigation. 60**:** 45-67.

Brooker, P. 2008. Air traffic safety: Continued evolution or a new paradigm? Aeronautical Journal. 112**:** 333-343.

Clemens, P. L. 2002. Fault tree analysis. JE Jacobs Severdurup.

Commission, N. R. 1975. Reactor safety study. An assessment of accident risks in US commercial nuclear power plants. Appendix XI. Analysis of comments on the draft WASH-1400 report. Nuclear Regulatory Commission, Washington, DC (USA).

Davis, D. 2005. SMC Systems Engineering Primer and Handbook. United States Air Force Space and Missile Systems Center.

Delaurentis, D. 2005. Understanding transportation as a system-of-systems design problem.

Delta Virtual Airlines. 2009. Boeing 747-400 Aircraft Operations Manual. First edition.

Dot, F. 2011. Introdcution to TCAS II Version 7.1.

Dunjó, J., Fthenakis, V., VÍlchez, J. A. and ArnaldoS, J. 2010. Hazard and operability (HAZOP) analysis. A literature review. Journal of hazardous materials. 173**:** 19-32.

Durga Rao, K., Gopika, V., Sanyasi rao, V. V. S., Kushwaha, H. S., Verma, A. K. and Srividya, A. 2009. Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. Reliability Engineering and System Safety. 94**:** 872-883.

Ericson, C. A. 2005. Event Tree Analysis. Hazard Analysis Techniques for System Safety. John Wiley and Sons, Inc.

Ericson, C. A. and LL, C. Fault tree analysis. System Safety Conference, Orlando, Florida, 1999: 1-9.

FAA 2011. NextGen Implementation Plan.

FAA 2015. FAA 7110 65.

Franklin, B. D., Shebl, N. A. and Barber, N. 2012. Failure mode and effects analysis: too little for too much? BMJ quality and safety. 21**:** 607-611.

Geisinger, K. 2003. Guide to methods and tools for safety analysis in air traffic management. Series Guide to Methods and Tools for Safety Analysis in Air Traffic Management. Global Aviation Information Network.

German Federal Bureau of Aircraft Accidents Investigation. 2002. Investigation Report.

Gilchrist, P. 1998. Boeing 747-400 (Airliner Color History). Osceola, WI: Motorbooks International.

Gran, B. A. and Helminen, A. 2001. A Bayesian belief network for reliability assessment. Computer Safety, Reliability and Security. 35-45.

Hall, A. D. and Fagen, R. E. 1956. Definition of system. General systems. 1**:** 18-28.

Hansman, R. J. and Odoni, A. 2009. Air traffic control. The Global Airline Industry**.** 377.

Haraldsdottir, A., Schwab, R. W. and Alcabin, M. S. 2001. Air traffic management capacity-driven operational concept through 2015. Progress in astronautics and aeronautics. 193**:** 9-26.

Harel, D. 1987. Statecharts: a visual formalism for complex systems. Science of Computer Programming. 8**:** 231-274.

Henley, E. J. and Kumamoto, H. 1985. Designing for reliability and safety control.

Jones, D. R. 2002. Runway incursion prevention system simulation evaluation. Digital Avionics Systems Conference, 2002. Proceedings. The 21st. 2: 11B4-1 - 11B4-12.

Kalman, R. 1959. On the general theory of control systems. Automatic Control, IRE Transactions on, 4: 481-492

Kenarangui, R. 1991. Event-tree analysis by fuzzy probability. Reliability, IEEE Transactions on. 40**:** 120-124.

Klamka, J. 2013. Controllability of dynamical systems. A survey. Bulletin of the Polish Academy of Sciences-Technical Sciences, 61**,** 335-342.

Klir, G. 1991. Facets of systems Science. IFSR International Series on Systems Science and Engineering (Vol. 7). Plenum Press, New York and London.

Knetcht, W. 1997. Developing a probabilistic metric of midair collision risk. Transportation Research Record: Journal of the Transportation Research Board**.** 26-32.

Kuchar, J. and Drumm, A. C. 2007. The traffic alert and collision avoidance system. Lincoln Laboratory Journal. 16**:** 277-295.

Landry, S. 2012. Intensity Control: A Concept for Automated Separation Assurance Safety and Function Allocation in NextGen. 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.

Leadbetter, D., Hussey, A., Lindsay, P., Neal, A. and Humphreys, M. 2001. Towards model based prediction of human error rates in interactive systems. Aust. Comput. Sci. Commun. 23： 42-49.

Lee, W.-S., Grosh, D. L., Tillman, F. A. and Lie, C. H. 1985. Fault Tree Analysis, Methods, and Applications A Review. Reliability, IEEE Transactions on. 34**:** 194-203.

Leveson, N. 2004. A new accident model for engineering safer systems. Safety Science. 42**:** 237-270.

Leveson, N. 2011. Engineering a safer world: Systems thinking applied to safety. MIT Press.

Leveson, N. 2015. A systems approach to risk management through leading safety indicators. Reliability Engineering and System Safety. 136**:** 17-34.

Leveson, N. G. and Stolzy, J. L. 1987. Safety analysis using Petri nets. IEEE Transactions on Software Engineering. 13**:** 386-397.

Liu, T. and Chiou, S. 1997. The application of Petri nets to failure analysis. Reliability Engineering and System Safety. 57**:** 129-142.

Mannan, S. 2012. Lees' Loss prevention in the process industries: Hazard identification, assessment and control, Butterworth-Heinemann.

Maurino, D. E., Reasonson, J., Johnstonton, N. and Lee, R. B. 1995. Beyond aviation human factors: Safety in high technology systems

Netjasov, F. 2010. Risk Analysis and Safety Assessment of Air Traffic Control System.

Netjasov, F., Vidosavljevic, A., Tosic, V., Everdij, M. H. C. and Blom, H. A. P. 2013. Development, validation and application of stochastically and dynamically coloured Petri net model of ACAS operations for safety assessment purposes. Transportation Research Part C: Emerging Technologies. 33**:** 167-195.

Nolan, M. 2010. Fundamentals of air traffic control, Cengage Learning

Oxford Dictionary, O. E. 1989. Oxford: Oxford university press.

ICAO. 2007. Manual on the Prevention of the Runway Incursions.

Potts, H. W., Anderson, J. E., Colligan, L., Leach, P., Davis, S. and Berman, J. 2014. Assessing the validity of prospective hazard analysis methods: a comparison of two techniques. BMC health services research. 14**:** 41.

Ptolemaeus, C. 2014. System Design, Modeling, and Simulation: Using Ptolemy II.

Rasmussen, J. 1997. Risk management in a dynamic society: A modelling problem. Safety Science. 27**:** 183-213.

Reason, J. 2000. Human error: models and management. Bmj. 320**:** 768-770.

Reynard, W. 1986. The development of the NASA aviation safety reporting system, National Aeronautics and Space Administration.

Rodgers, M. D., Mogford, R. H. and Strauch, B. 2000. Post hoc assessment of situation awareness in air traffic control incidents and major aircraft accidents. Situation awareness analysis and measurement**.** 73-112.

Saleh, J. H. and Bakolas, E. 2009. Augmenting the traditional defense-in-depth strategy with the concept of a diagnosable safety architecture. Reliability, Risk, and Safety, Three Volume Set. CRC Press.

Saleh, J. H., Marais, K. B., Bakolas, E. and Cowlagi, R. V. 2010. Highlights from the literature on accident causation and system safety: Review of major ideas, recent contributions, and challenges. Reliability Engineering and System Safety. 95**:** 1105-1116.

Savage, I. 2013. Comparing the fatality risks in United States transportation across modes and over time. Research in Transportation Economics. 43**:** 9-22.

Scovel III, C. L. and General, I. 2013. FAA's Progress and Challenges in Advancing the Next Generation Air Transportation System. Statement of the Honorable Calvin L Scovel III, Inspector General, US Department of Transportation before the Committee on Transportation and Infrastructure Subcommittee on Aviation United States House of Representatives, Washington DC, 17.

SESAR, J. U. 2012. European ATM Master Plan.

Shebl, N. A., Franklin, B. D. and Barber, N. 2009. Is failure mode and effect analysis reliable? Journal of patient safety. 5**:** 86-94.

Shorrock, S. T. 2005. Errors of memory in air traffic control. Safety Science. 43**:** 571-588.

Shorrock, S. T. 2007. Errors of perception in air traffic control. Safety science. 45**:** 890-904.

Shrivastava, S., Sonpar, K. and Pazzaglia, F. 2009. Normal accident theory versus high reliability theory: a resolution and call for an open systems view of accidents. Human relations. 62**:** 1357-1390.

Site, T. I. P. R. R. 1994. International Investigation Standards.

Siu, N. 1994. Risk assessment for dynamic systems: An overview. Reliability Engineering and System Safety. 43**:** 43-73.

SRC, E. S. R. C. 2005. EAM 2/GUI 5 Harmonisation of Safety Occurrence Severity and Risk Assessment.

Stamatis, D. H. 2003. Failure mode and effect analysis: FMEA from theory to execution. Asq Press.

Stroeve, S. H., Blom, H. A. P. and Bakker, G. J. 2009. Systemic accident risk assessment in air traffic by Monte Carlo simulation. Safety Science. 47**:** 238-249.

Stroeve, S. H., Blom, H. A. P. and Bakker, G. J. 2013. Contrasting safety assessments of a runway incursion scenario: Event sequence analysis versus multi-agent dynamic risk modelling. Reliability Engineering and System Safety. 109**:** 133-149.

Tanaka, H., Fan, L., Lai, F. and Toguchi, K. 1983. Fault-tree analysis by fuzzy probability. Reliability, IEEE Transactions On. 32**:** 453-457.

Skyttner, L. 2005. General systems theory: problems, perspectives, practice. World scientific.

Tian, J., Zhao, T. D. 2012. Controllability-involved Risk Assessment Model for Carrier-landing of Aircraft. 2012 Proceedings - Annual Reliability and Maintainability Symposium (Rams). 1-5.

The Boeing Company. 2007. 747 Specifications.

Torokhti, A., and Howlett, P. 1975. General systems theory: mathematical foundations. Academic Press.

Trucco, P., Cagno, E., Ruggeri, F. and Grande, O. 2008. A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation. Reliability Engineering and System Safety. 93**:** 845-856.

Varon, D. 2000. Air traffic control system. Google Patents.

Wasson, C. 2001. System Phases, Modes, and States: Solutions to Controversial Issues. Proceedings of the 21th Annual International Symposium of the International Council of Systems Engineering, 2011 Denver, Colorado.

Wickens, C. D. 1998. The future of air traffic control: Human operators and automation, National Academies Press.

Xu, X., Li, D. and Li, X. 2008. Research on safety assessment of flight separation. Hangkong Xuebao/Acta Aeronautica et Astronautica Sinica. 29**:** 1411-1418.

VITA

VITA

Jingjing Guo received her Bachelor of Engineering in Astronautics engineering from Harbin Institute of Technology in 2007. She was then promoted to the graduate school in the same program with the exemption of examination. She started her pursuit of doctoral degree in Aerospace Engineering at Purdue University in 2011, after a two year detour in Agricultural and Biological Engineering where she learned a great deal about hydraulics. Her current research interest is on the quantitative safety assessments of air traffic control systems from the perspective of system control and safety.